


```

RT mRNA,":
RL Submitted (FE8-2002) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: Phosphorylates uridine and cytidine to uridine
CC monophosphate and cytidine monophosphate. Does not phosphorylate
CC deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP
CC as a phosphate donor. Can also phosphorylate cytidine and uridine
CC nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-
CC thioridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-
CC benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-
CC methylcytidine, and N(4)-antioxylicytidine.
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- TISSUE SPECIFICITY: Ubiquitous.
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF237290; AAK8334.1; -
DR EMBL; AF254133; AAK49122.1; -
DR EMBL; AK022317; BAB14010.1; -
DR EMBL; AF125106; AAL75943.1; -
DR InterPro; IPR006083; PRK_UK.
DR InterPro; IPR00764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRFAMs; TIGR00235; udk; 1.
KM Transferrase; Kinase; ATP-binding.
FT NP_BIND 30 37 ATP (POTENTIAL).
FT CONFLICT 8 17 DCEBPAPAD -> GARARAGAN (IN REF. 4).
FT CONFLICT 56 57 QR -> HG (IN REF. 4).
FT CONFLICT 247 247 S -> T (IN REF. 4).
SQ SEQUENCE 277 AA; 31434 MW; AFD9BD92780CD502 CRC64;

Alignment Scores:
Pred. No.: 5,53e-88 Length: 277
Score: 1450.00 Matches: 277
Percent Similarity: 100.00% Conservative: 0
Best local Similarity: 100.00% Mismatches: 0
Query Match: 48.87% Gaps: 0
DB: 1

US-09-896-522-1 (1-1624) x UCK1_HUMAN (1-277)
QY 94 ATGGCTTCGGCGGAGGCGAAGATCGAGAGCCCCCGCGGAGCGGACCGTCCGCAC 153
DB 1 MetalaetraiaiglygluapcysgluSerProalaprogluAlaaspargProhis 20
QY 154 CAGGCGCCCTTCTGATAGGGGTGAGCGCGGCGACTGCGACGCGGAAGTGCAGCGTGT 213
DB 21 GlnargProheliuegllyvalSercllylthralaserglylserThvalCys 40
QY 214 GAGAGATCATGAGTGTCTGGAGCAGAACGAGGTGGAACGCGGACGCGAGGTGTC 273
DB 41 GluylsileMetGluLeuenglGlnabnsluValGluGlnaGlnaGlyValVal 60
QY 274 ATCTGAGCCGAGCAGAGTTTACAAGTCTTGACCGCAGAGCAGAGCCGAGCCTTG 333
DB 61 IleleuSerGlnaPargPheTylylValLeuThrAlaGluGlnyAlaValLeu 80
QY 334 AAAGAGATCAATTTTGACATCCAGATCCCTTTGATATGATTTGATGACAGACT 393
DB 81 LyselylGlnTyraenPheAspHisProkspAlaPheAspAlaPheMeHisArgThr 100
QY 394 CTGAAGACATCGTGAGAGGGAACCGGTGAGGTGCCAGCTTATGATTTTGACACAC 453
DB 101 LeuylAsnIleValGluGlylValGluValProThrTyraPheValThHis 120

```

```

QY 454 TCAAGGTTACCAAGACCAACGCGTGTCTAACCCCTGCGGAGGTGGTGTCTTGAGGGCATC 513
DB 121 SerArgleuProdluThrThrValValTyProAlaAspValValLeuPheGluGlylle 140
QY 514 TTGGTGTTCACAGCCAGAGAGATCCCGGACATGTTCCACCTGCGCCTTCTTGACAC 573
DB 141 LeuValPheylSerGlnGluIleArgAspMetPheHisIleuArgleuPheValAspThr 160
QY 574 GACTCCGAGCTCAGGCTGTCTGCAAGAGTTCTCCGGAGCTGGCGGAGGAGACCTG 633
DB 161 AspSerAspValArgleuSerArgValValLeuArgAspValArgValArgGlyArgAspLeu 180
QY 634 GACACATTCGACGAGATACACACCTTCGTGAACCGGCGCTTGAGAGGTCTGCTCG 693
DB 181 GluGlnIleuThrGlnTyraenPheValThrPheValLysProAlaPheGluGluPheCysLeu 200
QY 694 CCGACAAGAGATGATCCGATGATCATCCACAGAGAGTGAGACAATATGTTGCCATC 753
DB 201 ProThrLysylSerTyraenPheValIleleleProArgGlyValAspAsnMetValAlaIle 220
QY 754 AACCTGATCGTGCAGACATCCAGACATTTGTAATGTGTGACATCTTGCAATGCAACCGA 813
DB 221 AsnIleuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIlystrPHisArg 240
QY 814 GAGAGGTTCGAATGGGCGGAGCTCAAGCGGACCTTTTCGAGCGAGGGGACCGCTGG 873
DB 241 GlyIylSerAsnIlylArgserTylylAspArgThrPheSerGluProdluAspHisProGly 260
QY 874 ATCTGATCCTCTTGCAACCGTGCATTTGATGATCCAGCAGACAGACCCGAC 924
DB 261 MetleuThrSerIlylAspArgserHisIleuGluIleSerSerArgProhis 277

RESULT 2
UCK1_MOUSE
ID UCK1_MOUSE STANDARD; PRT; 277 AA.
AC P52623;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Uridine-cytidine kinase 1 (EC 2.7.1.48) (UCK 1) (Uridine
DE monophosphokinase 1) (Cytidine monophosphokinase 1).
GN UCK1 OR UMPK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ueda T.B., Tohyuki S., Carninci P., Frange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Wallaby S.J.,
RA Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Munzy D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE OF 18-277 FROM N.A.

```

RC	TISSUE=Brain;	
RX	MEDLINE=97108719; PubMed=8951040;	
RA	Ropp P.A., Traut T.W.;	
RT	"Cloning and expression of a cDNA encoding uridine kinase from mouse	
RL	brain.";	
RL	Arch. Biochem. Biophys. 336:105-112(1996).	
CC	-1- FUNCTION: Phosphorylates uridine and cytidine to uridine	
CC	monophosphate and cytidine monophosphate. Does not phosphorylate	
CC	deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP	
CC	as a phosphate donor. Can also phosphorylate cytidine and uridine	
CC	nucleoside analogs such as 6-thauridine, 5-fluorouridine, 4-	
CC	thiouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-	
CC	benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-	
CC	-1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.	
CC	-1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.	
CC	-1- PATHWAY: Pyrimidine salvage pathway.	
CC	-1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on ways	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (see http://www.isb-sib.ch/announce/	
CC	or send an email to license@ib-sib.ch).	
CC	-----	
DR	EMBL; BC025146; AAH25146.1; -.	
DR	EMBL; L31783; AAB50568.1; -.	
DR	MGI; MGI:98904; Umpk.	
DR	InterPro: IPR006083; PRK URK.	
DR	InterPro: IPR000764; Uridine_kin.	
DR	Pfam; PF00485; PRK; 1.	
DR	PRINTS; PR00988; URIDINKINASE.	
DR	TIGRFAMs; TIGR00235; udk; 1.	
KW	Transferase; Kinase; ATP-binding.	
FT	NP BIND 30 37	
FT	SEQUENCE 277 AA; 31068 MW; 3EBB3C4187FAEB4A CRC64;	

[illegible]

Oy	451	TCAAAGTTACCGAGACA	CCGGTGGTCA	ACCCTCGGAC	GGTGGTCTCGTTGAGGGCATC	513
Dd	121	SeTArGLeuPrCoGluNhrThrValValIyTrProAlaSerValValIeuphneGluGlyIle				140
Oy	514	TTGGTTTCTACAGCCAGCAGAGATCCGGGACATGTATTCACCTGGCCCTCTTGGTGACACC				573
Dd	141	LeuValPheTyIThrGlnGlnIleLeuArgSerPheNhsIeuAlaGleuPheValAspThr				160
Oy	574	GATCTCCGAGCTCAAGGCTGTCTTGAAAGATTCTCCGGACGTGGCCGACGGAGGGACCTG				633
Dd	161	AspSerAspValAlaGleuSerArGArgValIeuaTrgAspValGlnAlaGgGlyAArgAspLeu				180
Oy	634	GAGCAGATTCTGACCGAGTACACCACTCTCGGAAAGCCGGGCTTCAGAGAGTTCTGGCTG				693
Dd	181	GIuGlnIleuThrGlnIlyrThrAlaPheValIlySerProAlaPheGlnGluPheCysLeu				200
Oy	694	CCGACAAAGATATGCCGATGTGCATCCACGAGAGTGCACAAATATGGTTGCCATC				753
Dd	201	ProThrIlyLysTyIAlaAspValIleIleProArgIlyValAlaAspMetValAlaIle				220
Oy	754	AACCTGATGTGTCACACATCCAGGACATTCGTGAATGTGTCACATCTCCAAATGGCACCGA				813
Dd	221	AsnIeuIleValGlnHisIleIleAspIleIeuSerGlnGlySerIeuCysIlyAspGlnAsArg				240
Oy	814	GGAGGGTCCAAATGGGCGGAGTACCAAGCGGACCTTTTCTGAGCCGAGGGGACCAACCTGGG				873
Dd	241	GIyGlyProAsnGlyArgAsnHisIlyAspArgThrPheProGluProGlyIleAspHisProGly				260
Oy	874	ATGCTGACCTCTGGGAAACGGTCACATTTGGAGATCCAGACGACCAACCCGAC				924
Dd	261	ValIeuAlaThrGlyLysArgSerHisIleuGlnSerIeuSerIeuArgProHis				277

ID	UICK2_HUMAN	STANDARD;	PRT;	261 AA.
AC	Q9BZX2; Q96KG5;			
DT	28-FEB-2003 (Rel. 41, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Uridine-cytidine kinase 2 (EC 2.7.1.48) (UICK 2) (Uridine monophosphokinase 2) (Cytidine monophosphokinase 2).			
GN	UICK2 OR UMPK.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxId=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND CHARACTERIZATION.			
RX	MEDLINE=21203813; PubMed=11306702;			
RA	Van Rompay A.R., Norda A., Linden K., Johansson M., Karlsson A.;			
RT	"Phosphorylation of uridine and cytidine nucleoside analogs by two human uridine-cytidine kinases.";			
RL	Mol. Pharmacol. 59:1181-1186(2001).			
RN	[2]			
RP	SEQUENCE OF 15-261 FROM N.A.			
RC	TISSUE=Fibrosarcoma;			
RX	MEDLINE=21385121; PubMed=11494055;			
RA	Koizumi K., Shimanoto Y., Azuma A., Wataya Y., Matsuda A., Sasaki T.,			
RA	Fukushima M.;			
RT	"Cloning and expression of uridine/cytidine kinase cDNA from human fibrosarcoma cells.";			
RL	Int. J. Mol. Med. 8:273-278(2001).			
CC	-1- FUNCTION: Phosphorylates uridine and cytidine to uridine monophosphate and cytidine monophosphate. Does not phosphorylate deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP as a phosphate donor. Can also phosphorylate cytidine and uridine nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-chlorouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-methylcytidine, and N(4)-anisoylcytidine.			
CC	-1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.			
CC	-1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.			
CC	-1- PATHWAY: Pyrimidine salvage pathway.			

Db	184	GlnTrlrlernhrPheVallyspRoalAphEglnUlnubheCysIseuProThrllyLyvetyr	203
Qy	709	GCCGATGTCATCATCCACGAGAGTGGACACATATGTTGGTCATCAACCTGATGTGCAG	768
Db	204	AlasPValIlellePProArgGlyAlaAspAsnIeuValIleValGln	223
Qy	769	CACATCCAGGACATTTCGATTCGGTGGACATCTGCAAATGACACCGAGAGGATCCAAATGGC	828
Db	224	HisIleGlnAspIleLeuAsnGly	232
Qy	829	CGAGCTACAAAGCGGACCTTTTCTGAGCCAGGGACACCTTGAGGCTGACCTTGAC	888
Db	233	ProSerLysArgGlnThr	249
Qy	889	AAAGGTCACATTTGGAGTCCAGACAGACACCCAC	924
Db	250	LysArgGlnAlaSerGlnSerSerSerArgProHis	261

RESULT 4

UCLK2_MOUSE

ID	UCLK2_MOUSE	STANDARD:	PRT:	261 AA.
AC	Q99PM9			
DT	28-FEB-2003	(Rel. 41, Created)		
DT	28-FEB-2003	(Rel. 41, Last sequence update)		
DT	28-FEB-2003	(Rel. 41, Last annotation update)		
DE	Uridine-cytidine kinase 2 (EC 2.7.1.48) (UCLK 2) (Uridine monophosphokinase 2) (Cytidine monophosphokinase 2) .			
GN	UCLK2 OR UMPK.			
OS	Mus musculus (Mouse) .			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21203813; Pubmed=11306702;			
RA	Van Rompay A.R., Norda A., Lindén K., Johansson M., Karlsson A.;			
RT	"Phosphorylation of uridine and cytidine nucleoside analogs by two human uridine-cytidine kinases."			
RL	Mol. Pharmacol. 59:1181-1186(2001).			
CC	-1- FUNCTION: Phosphorylates uridine and cytidine to uridine monophosphate and cytidine monophosphate. Does not phosphorylate deoxyribonucleosides or purine ribonucleosides. Can use ATP or GTP as a phosphate donor. Can also phosphorylate cytidine and uridine nucleoside analogs such as 6-azauridine, 5-fluorouridine, 4-thiouridine, 5-bromouridine, N(4)-acetylcytidine, N(4)-benzoylcytidine, 5-fluorocytidine, 2-thiocytidine, 5-methylcytidine, and N(4)-antisoacylcytidine (By similarity) .			
CC	-1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.			
CC	-1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.			
CC	-1- PATHWAY: Pyrimidine salvage pathway.			
CC	-1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb.ch/announce/> or send an email to license@isb-sib.ch).

CC MDL; AF236636; UAK14052.1; . | | | || DR | MGD; MGI:1931744; UCLK2. | | | |
DR	InterPro; IPR006082; PRK.			
DR	InterPro; IPR006083; PRK_URK.			
DR	InterPro; IPR000764; Uridine_kin.			
DR	Pfam; PF00485; PRK.1.			
DR	PRINTS; PR00478; PHRIIDKINASE.			
DR	PRINTS; PR00989; URIDINKINASE.			
DR	TIGRFAMs; TIGR00235; udk.1.			
TM	Transferase; kinase; ATP-binding.			
NP_BIND	27	ATP (POTENTIAL).		
SEQUENCE	261 AA; 29404 MW; 780AA3CFE5CA8153 CRC64;			

Alignment Scores:

Pred. No.: 3,39e-41 Length: 260
 Score: 738.00 Matches: 143
 Percent Similarity: 82.94% Conservative: 32
 Best Local Similarity: 67.77% Mismatches: 34
 Query Match: 24.87% Indels: 2
 Gaps: 2

US-09-896-522-1 (1-1624) x UCK_DROME (1-260)

QY 160 CCCTTCCTGATAGGAGGTAGCGGGGCACTGCCACCGGGAAGTGCACCGCTGTGTAGAG 219
 Db 27 Prothelutleuilegylvalalaglylthralserglylservthvalcyslyls 46
 QY 220 ATCATGAGTGTCTGGAGCAAGACGAGGTGGAACGCGGAGCGGAGGTGATCTGCTG 279
 Db 47 Illemtuglunleuglylgnalaglumetaphsthringlnaglnvalvalserlle 66
 QY 280 AGCAGAGCAGGTTCTTACAGAGTCTCGACGAGCAAGAGCCAGGCTTGAAGA 339
 Db 67 SerghnaSerPheTyArgGluLeuthrProalaglnulysalvalaglnlygily 86
 QY 340 CAGTACATTTTGCATCCAGATCCCTTGTATATGATTGTGACACAGACTCAAG 399
 Db 87 LeuthenPheasphtsproaspalaheasnuglnleuMetTySerThrlleugln 106
 QY 400 AACATCGGAGGCGCAAAACGAGGAGGTGCGGACCTTGTATTTGACACACTCA-- 456
 Db 107 AsnleleuylsglyhilelylvalglnleproserTyArgThrsAsnleu 126
 QY 457 AGGTTACAGAGACACCGTGTCTTACCTCGGACGCTGTCTTGTGAGGCACTTG 516
 Db 127 AsphegluanValleuValilleTyxProalaspValalleuPhegluglylleu 146
 QY 517 GTGTTTACACCGAGAGATCCGGGACATGTTCCACCTCGGCGCTTCTGTGACACGAC 576
 Db 147 ValPheTyRheProlyslargluueuPhehshetlyslleuPheValAspThrasp 166
 QY 577 TCCGACGTGAGTGTCTCGAAGATTCTCCGGGACGTG--CGCCGAGGAGGAGACTG 633
 Db 167 SeraphthraArgleuAlaArgValProArgAspIleasnuglnargglyArgAspleu 186
 QY 634 GAGCAGATTGACGACGACTACCACTTGTGTAAGCCGCGCTTGCAGAGACTTGCCTG 693
 Db 187 AspAlaValleuThrgIntyMetThrpheVallyspProalasphegluglnPheCysSer 206
 QY 694 CCGACAAAGAGTATGCCAGATGATATCCACAGAGAGGACATATGTTGCCATC 753
 Db 207 ProthrylsylsPheAlaAspValilleleProArgglyAlaAspAsnthValAlaile 226
 QY 754 AACCTGATCGTGCAGACATCCAGACATCTTG 786
 Db 227 Aspleuilevalhshisileglylglulleu 237

RESULT 6
 URL_HUMAN STANDARD; PRT; 548 AA.
 AC Q9NMZ5; O9H322;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Uridine Kinase-like 1.
 GN URKLI.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Matanabe K., Kumagai A., Itakura S., Yamazaki M., Taahiro H., Ota T.,
 RA Suzuki Y., Ohayashi M., Nishi T., Shibahara T., Tanaka T.,
 RA Nakamura Y., Isoigai T., Sugano S.;
 RT "NEDO human cDNA sequencing project."

RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Bagguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Copley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Graffam D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J.L., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Lehesvahti M.H., Leverhna M.A., Lloyd C., Lloyd D.M., Lovell J.D.,
 RA Marsh V.L., Martin S.L., McCormachie L.J., McElay K., McMuray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillips B.J.C.T., Prachinang S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showreen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Sycamore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitteker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilmshurst L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RT "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AK000524; BA91230.1; -;
 DR EMBL; AL118506; CAC15497.1; -;
 DR HSSP; Q26998; 1BD3.
 DR Genew; HGNC:15938; URKLI.
 DR InterPro; IPR006082; PRK.
 DR InterPro; IPR006083; PRK.
 DR InterPro; IPR00764; Uridine_kin.
 DR Pfam; PF00485; PRK.1.
 DR PRINTS; PR00478; PHRIBLKINASE.
 DR PRINTS; PR00988; URIDINKINASE.
 DR TIGRFAMs; TIGR00235; udk.1.
 KW Transferase; Kinase; ATP-binding.
 FT NP_BIND 105 112 ATP (POTENTIAL).
 FT CONFLICT 219 228 MISSING (IN REF. 2).
 FT CONFLICT 290 290 N -> D (IN REF. 2).
 SQ SEQUENCE 548 AA; 61139 MW; 0CD03697E02FE7DB CRC64;

Alignment Scores:
 Pred. No.: 7.14e-29 Length: 548
 Score: 551.50 Matches: 122
 Percent Similarity: 55.37% Conservative: 48
 Best Local Similarity: 39.74% Mismatches: 68
 Query Match: 18.59% Indels: 69
 Gaps: 8

US-09-896-522-1 (1-1624) x URL_HUMAN (1-548)
 QY 7 TCGCTCCGACCTCGGCGCTGAGCGGCGCGGCGGAGAGGCGGCGGCGG 66
 Db 16 SerPheProhrrala-----ArgAspThrProGlyArg----- 26
 QY 67 ACCGATGCGCGGAGAGGAGCGGAGCGGAGATGCTTGGCGGAGCGAGAGAC----- 117


```

Db 27 -----GlnAlaGluLysSerLurhAlaCysGluAaPaGSerAsn 40
Qy 117 -----
Db 41 AlaGluSerLeuPaArgLeuProProValGlyThrGlyArgSerProArgLysArg 60
Qy 118 -----TGCGAGAGCGCCCGCGCGCG----- 135
Db 61 ThrThrSerGlnCysLysSerGluProProLeuLeuArgThrSerLysArgThrLleTyr 80
Qy 136 GAGCGCCAGCCGCTCG-----CACGAGCGCCCTTCTCG 168
Db 81 ThrAlaGlyArgProProTyrPyrAsnGluHisGlyThrGlnSerLysGluAlaPheAla 100
Qy 169 ATAGGGGTAGCGCGCGCGCGCGCGCGAGAGTGCACCGTGTGTGAAGATCATGAG 228
Db 101 IleGlyLysGlyGlySerAlaSerGlyLysThrThrValAlaArgMetIleIleGlu 120
Qy 229 TTGCTGGAGAGAAACGAGTGGAGAACAGCGGAGGAGGTGTCATCTTCAGCGAGAC 288
Db 121 AlaLeu-----AspValProTyr-----ValValLeuLeuSerMetAsp 133
Qy 289 AGGTTCTACAGAGCTCTGAGCGAGAGAGAGAGCGCGCGCGCGCGCGCGCGCGCG 348
Db 134 SerPheTyrLysValLeuThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 153
Qy 349 TTGACCATCCAGATGCTGTTGATGATGATGATGATGATGATGATGATGATGATGATG 408
Db 154 PheAspHisProAspAlaPheAspPheAspLeuLleSerThrLeuLysLysLys 173
Qy 409 GAGGGGAAACGGTGGAGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 468
Db 174 GlnGlyLysSerValLysValProLleTyrAspPheThrThrHisSerArgLysLysAsp 193
Qy 469 ACCAGCGTGTCTACCTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 528
Db 194 TrpLysThrLeuTyrGlyAlaAsnValIleIlePheGlnGlyIleMetAlaPheAlaAsp 213
Qy 529 CAGAGATCCGCGAGCATGTTCCATCGCGCGCTCTTCGTGACACCGAGCTCCGAGCG 588
Db 214 LysThrLeuLeuGlnLeuLeuAspMetLysIlePheValAspThrAspSerAspLleArg 223
Qy 589 CTGTCTCGAGAGTTCCTCCGCGAGCGT---CGCGAGAGAGAGCGCGCGCGCGAGTCTG 645
Db 234 LeuValArgArgLeuArgArgAspLleSerGluArgGlyArgAspLleGlnGlyValIle 253
Qy 646 ACCGAGTACACCACTCTGAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAG 705
Db 254 LysGlnTyrAsnLysPheValLysProSerPheAspGlnTyrIleGlnProThrMetArg 273
Qy 706 TATGCCGATGTGATCATCCACGAGAGAGTGAACAATATGTGTCCATCAACTGTATG 765
Db 274 LeuAlaAspLleValValProArgGlySerGlyAsnThrValAlaIleAsnLeuLleVal 293
Qy 766 CAGCAGATCCAGAGATTCG 786
Db 294 GlnHisValHisSerGlnLeu 300

```

```

RESULT 7
URL_MOUSE
ID URL_MOUSE STANDARD; PRT, 548 AA.
AC 091YF3,
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Uridine kinase-like 1.
GN URK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932;
RA Strasserberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Stemann C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaefer T.E.,
RA Brownstein M.J., Ueda T.B., Toshlyuk S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Maira M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -i- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL, BC016535; AH16535.1; --
DR MGD; MGI:1915806; 1110007H10R1X.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK_1.
DR PRINTS; PR00478; PHR1BKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TRIPRAMS; TRIGR00235; udk. 1.
DR Transferase; Kinase; ATP-binding.
FT NP BIND 105 112 ATP (POTENTIAL).
SQ SEQUENCE 548 AA; 60841 MW; 99A5CA23EB9B8525 CRC64;

```

```

Alignment Scores:
Pred. No.: 3,01e-28 Length: 548
Score: 542.00 Matches: 116
Percent Similarity: 62.55% Conservative: 46
Best Local Similarity: 44.79% Mismatch: 79
Query Match: 18.27% Indels: 19
DB: 1 Gaps: 5

```

```

US-09-896-522-1 (1-1624) x URL_MOUSE (1-548)
Qy 44 CCGGGAGAGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 88
Db 49 ProProValGlyThrGlnSerLysArgProArgLysArgThrThrSerGlnCysArgLys 68
Qy 89 CCGAGATGCTTCGCGCGGAGAGGAGAGAGTGCAGAGCGCGCGCGCGCGCGCGCGCG 148
Db 69 ProProLeuLeuArgThrSerLysArgThrLleTyrThrAlaGlyArgProProTyr 88
Qy 149 CGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 192
Db 89 AenGluHisGlyThrGlnSerLysGluAlaPheAlaIleGlyLeuGlyGlySerAla 108
Qy 193 AGCGGAGATCCAGCGCGTGTGTGATGAGAGATGATGATGATGATGATGATGATGATG 252
Db 109 SerGlyLysThrThrValAlaArgMetIleIleGluAlaLeu-----AspValPro 125
Qy 253 CAGCGGAGCGGAGAGTGTCTCATCTTCGAGCGAGAGAGTTCATCAAGTCTCGAGCGCA 312

```

```

Db      1266 Trp-----ValValIleuenseMetAspSerPheTyrIlyValIleuThrGln 144
QY      313 GAGCGAAGAGCCCAAGCCCTGTGAAAGACAGTACATTTTGGACATCCAGATGCCCTTGTAT 372
      ::|||:::  ::|||:::
Db      142 GInGInGInGInGInGInAlaIaIaCysAsnAsnPheAsnPhaSPHisIeTroAspAlaPheAsp 161
QY      373 AATGATTTGATGTCACAGACACTCTGAAAGAACATCGGAGAGGCGAAGAAACGGTGGAGGCGCG 432
      |||||:::  |||||:::  ::|||:::  ::|||:::  ::|||:::
Db      162 PheAspLeuIleIleIeSerThrIleuIlySleuIySGInGlyArgSerValGlnValPro 181
QY      433 ACCTTATGATTTTGTGACACACACTCAGAGTTTACCAAGAGACCAAGGTGGTCTTACCTGGCGAC 492
      |||||:::  |||||:::  ::|||:::  ::|||:::
Db      182 IleTyrAspPheThrThrHisSerArgIlySlyAspTrpIlySerThrIleuTyrGlyAlaAsn 201
QY      493 GTGGTTCGTGTTGAGAGCACTTCTGGGTCTACAGCAGAGATCCGAGACAGTGTCCAC 552
      |||||:::  |||||:::  |||||:::  ::|||:::
Db      202 ValIleIleIlePheGInGlyIleMetAlaPheAlaAspIlySThrIleuGInIleuIleuAsp 221
QY      553 CTGCGCCTCTTGTGTGACACCGACTCCGACGTCAAGGCTGTCTGAAAGTGTCTCCGGAC 612
      |||||:::  |||||:::  |||||:::  ::|||:::
Db      222 MetIlyIlePheValAspThrAspSerAspIleArgLeuValArgArgIleuArgArgAsp 241
QY      613 GTG---CCCGGAGGAGGAGCACTCGAGAGCATTCGTGACCGCAAGTACACACCTTCGTGAAG 669
      ::|||:::  |||||:::  |||||:::  ::|||:::
Db      242 IleSerGluArgGlyArgAspIleGInGlyValIleIySGInTyrAsnIlyAspIheValIyS 261
QY      670 CCGGCGCTTCGAGAGATTCCTGCCTCGCGACCAAGAGTAGTACCGATGTATCATCCACAG 729
      |||||:::  |||||:::  |||||:::  ::|||:::
Db      262 ProAlaPheAspGInTyrIleGInProThrMetArgLeuAlaAspIleValAlaProArg 281
QY      730 GGAGTGGACATATATGTTGCCATCAACTGATCGTGGACAGACATCCAGACATTCTG 786
      |||||:::  |||||:::  |||||:::  ::|||:::
Db      282 GlySerGlyAsnThrValAlaIleAspIleuIleValGlnHisValHisSerGlnIleu 300

RESULT 8
ID      UCK CAEEL          STANDARD;          PRT;          229 AA.
UC      UCK CAEEL
AC      Q17413;
DT      28-FEB-2003 (Rel. 41, Created)
DT      28-FEB-2003 (Rel. 41, Last sequence update)
DT      28-FEB-2003 (Rel. 41, Last annotation update)
DE      Probable uridine-cytidine kinase (EC 2.7.1.48) (UCK) (Uridine
DE      monophosphokinase) (Cytidine monophosphokinase).
GN      B0001.4.
OS      Caenorhabditis elegans.
OC      Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC      Rhabditidae; Pelodexinae; Caenorhabditis.
OX      NCBI_Taxid=6239;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=Bristol NZ;
RA      Sime M.;
RL      Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
CC      -1 CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC      -1 CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC      -1 PATHWAY: Pyrimidine salvage pathway.
CC      -1 SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sib.ch/announce
CC      or send an email to license@isb-sib.ch).
CC      -----
DR      EMBL; Z69634; CAA93453.1; -
DR      PIR; T18629; T18629.
DR      Wormpep; B0001.4; CE05142.
DR      InterPro; IPR006082; PRK.
DR      InterPro; IPR006083; PRK_UK.
DR      InterPro; IPR000764; Uridine_kin.
DR      Pfam; PF00485; PRK; 1.

```

[illegible]


```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Maeni N.,
RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -|- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -|- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -|- PATHWAY: Pyrimidine salvage pathway.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AP001511; BAB04394.1; -
DR PIR; C83809; C83809.
DR HAMAP; MF_00551; - 1.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00478; PHRIBLKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRPFAMs; TIGR00235; udk; 1.
DR Transferase; Kinase; ATP-binding; Complete proteome.
DR NP BIND . 12 19 ATP (POTENTIAL).
SQ SEQUENCE 211 AA; 24387 MW; C2AFR2CB0030520B CRC64;

Alignment Scores:
Pred. No.: 8,48e-24 Length: 211
Score: 473.00 Matches: 93
Percent Similarity: 66.51% Conservative: 50
Best Local Similarity: 43.26% Mismatches: 56
Query Match: 15.94% Indels: 16
DB: Gaps: 4

US-09-896-522-1 (1-1624) x URK_BACHD (1-211)
QY 154 CAGCGCCCTTCCTGATAGGGGAGGCGGCACTGCAGCGGAAGTGCAGCGTGT 213
DB 3 LysAspProLeuIleIleGlyValAlaGlyGlyThrGlySerGlyLysThrThrValAla 22
QY 214 GAGAGATCATGAGAGTTGCTGGAGACAGAGATGAGAGCGGCGAGCGGAAGTGTG 273
DB 23 LysGluIleIlehe-----TyrGlnhehnsnGluLysSerIleVal 35
QY 274 ATCTTAGAGCGAGACAGTTCTTCAAG-----GTCTTAGCGGACAGCAAGAGCC 324
DB 36 LeuIleGlnIleAspAlaTyrTyrLysAspGlnSerGlnIleuSerLeuGlnGluArgLeu 55
QY 325 AAGCGCTTGAAGAGCATGATTTTTCAGATCCAGATCCAGTTCATATATGATTTGAG 384
DB 56 -----GlnThrAsnTyrAspHisProLeuAlaIleheAspAsnLeuLeu 70
QY 385 CACAGAGCTCTGAAGACATCTGTGAGGAGGAGAAACGGTGGAGGAGTGCAGCTATGATTT 444
DB 71 IleGluHisLeuHisSerLeuLeuAanglyGlnAlaIleGluLysProValTyrAspTyr 90
QY 445 GTACACACCTCAAGGTTTACAGAGACACCGGTGTTTACCTTCGGAAGCTGTTGTTT 504
DB 91 LysLeuHisThrArgSerAanglyValIleLeuValGluProLysAspValIleIleLeu 110

```

```

QY 505 GAGGACATCTGTGTTCTTACAGCCAGAGATCCGGAGCATGTTCCAGCGGCTCTTC 564
DB 111 GluGlyIleLeuLeuLeuGluAspGluArgLeuArgGluLeuMetCysPheLeuPhe 130
QY 565 GTGACACCGACTCCAGCGTCCAGCTGTCTCGAAGAGTTCTCCGGAGCGTGC---CGA 621
DB 131 ValAspThrAspAlaAspIleArgIleIleArgArgMetValArgAspIleArgGluArg 150
QY 622 GGGAGGACCTGGAGAGATTTCTGACGCGATACACCACTTGTGTGAAGCGGCTTCGAG 681
DB 151 GlyArgThrLeuGluSerValIleGluGlnTyrThrLysValValArgProMetHisMet 170
QY 682 GAGTTCGCTGCGCGCAAGAAAGTATGCGGATGATCATCCAGACGAGATGCAAT 741
DB 171 GlnPheIleGluProThrLysArgTyrAlaAspValIleIleProGluGlyGlnAsn 190
QY 742 ATGGTTCATCAACCTGATCGTGCAGACATCAAGACATTCGTG 786
DB 191 ArgValAlaIleAspLeuMetValThrLysIleArgAlaIleIle 205

RESULT 10
URK_CLOPE
ID URK_CLOPE STANDARD; PRT; 208 AA.
AC Q8XJ16;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine
DE monophosphokinase).
GN UDK OR CPE1770.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX MEDLINE=21664373; PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Oshima K., Yamashita A.,
RA Shiba T., Ogasawara N., Hatiori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
CC -|- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -|- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -|- PATHWAY: Pyrimidine salvage pathway.
CC -|- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -|- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AP003191; BAB1476.1; -
DR HAMAP; MF_00551; - 1.
DR InterPro; IPR006082; PRK.
DR InterPro; IPR006083; PRK_URK.
DR InterPro; IPR000764; Uridine_kin.
DR Pfam; PF00485; PRK; 1.
DR PRINTS; PR00478; PHRIBLKINASE.
DR PRINTS; PR00988; URIDINKINASE.
DR TIGRPFAMs; TIGR00235; udk; 1.
DR Transferase; Kinase; ATP-binding; Complete proteome.
DR NP BIND . 18 23884 MW; 941BC368740DDE5E CRC64;
SQ SEQUENCE 208 AA; 23884 MW; 941BC368740DDE5E CRC64;

Alignment Scores:
Pred. No.: 1.8e-23 Length: 208
Score: 468.00 Matches: 88

```

Percent Similarity:	66.51%	Conservative:	55
Best Local Similarity:	40.93%	Mismatches:	56
Query Match:	15.77%	Indels:	16
DB:	1	Gaps:	4
US-09-896-522-1 (-1-1624) x URK_CLOPE (1-1208)			
QY	154 CAGCGGCGCTTCTGTATAGGGGTGAGCGCGTCCGACCTGCCAGCGGGAAGTGTGACCGGTGT	213	
DB	2 LysArgProIleAerPheIleGlyIleThrGlyGlyThrGlySerGlyLysSerThrIleLea	21	
QY	214 GAGAAAGATATGAGAGTGTGCTGGGACAGAAACGAGGTGGAAACAGGGGACCGGAAGTGC	273	
DB	22 LysGluIleTyArgGlnPheGlyGluAsp-----CysIleAla	34	
QY	274 ATCTGAGCCGAGACAGAGTTCTTAACA-----GTCTGACGGCAGACGAGAAGGCC	324	
DB	35 MetIleGlnGlnAspSerTyTrpLysAspGlnSerHISLeuSerMetGlnAspArgVal	54	
QY	325 AAGGCGCTTAAGACAGTCAATTTTGACATCCAGATCCGATCCCTTGATTAATGATTGATG	384	
DB	55 LysThr-----AsnTyAspHisProAsnAlaPheAspAsnAlaLeu	69	
QY	385 CACAGGACTGTGAAGAATCTCGAGAGGGCAAAACGGTGAAGGTGGCGACTATGATTT	444	
DB	70 ValSerHisLeuGlnSerLeuLeuAsnGlyHisSerIleGlnLysProSerTyAspPhe	89	
QY	445 GTGACACACTCAAGGTATTACACAGACACACCGATGATCCCTCGGACGTGTCTGTT	504	
DB	90 SerIleHisAsnArgIleGluAspThrThrLysValGlnProLysGluIleValIleVal	109	
QY	505 GAGGCGACTTGATGTTCTTAACAGCCAGAGATCCGGAGCATGTTTCCACTGCGCTTTC	564	
DB	110 GlnGlyIleLeuIleLeuGlnLysProArgIleArgGlnLeuLeuAspIleLysIleTy	129	
QY	555 GTGACACCGGACTCCGACGTCGCTGTCTCGAAGACTTCCGGGACGCG---CGCGCA	621	
DB	130 ValAspThrAspAlaAspValArgIleIleArgArgMetValArgAspIleAsnGluArg	149	
QY	622 GGGAGGAGCATCGAGAGATTTGACGACGATACACCACTTCCTGAAGCGCGGCTTCGAG	681	
DB	150 GlyAlaGlnThrMetGlnSerValIleAsnGlnTyLeuAsnValValLysProMetHisAsn	159	
QY	682 GAGTTCTGCTCCGCAAAAGATATGCCGATGTATCATCCACAGAGAGTGGACAAT	741	
DB	170 GlnPheThrGlnProThrLysLysPheAlaAspIleIleProGlnGlyGlyHisAsn	189	
QY	742 ATGTTTCCATCACTGATTCGGACGACATCCAGACATTTCTG	786	
DB	190 LysValAlaIleAspIleIleValAlaLysIleLysGluValLeu	204	
RESULT 11			
URK_LACUA STANDARD; PRT; 206 AA.			
AC	Q9CE21;		
DT	28-FEB-2003 (Rel. 41, Created)		
DT	28-FEB-2003 (Rel. 41, Last sequence update)		
DT	28-FEB-2003 (Rel. 41, Last annotation update)		
DE	uridine kinase (EC 2.7.1.48) (uridine monophosphokinase) (Cytidine monophosphokinase).		
GN	UDK OR L16160.		
OS	Lactococcus lactis (subsp. lactis) (Streptococcus lactis).		
OC	Bacteria; Firmicutes; Lactobacillales; Streptococcaceae; Lactococcus.		
OX	NCBI_TaxID=1350;		
RN	SEQUENCE FROM N.A.		
RC	STRAIN=IL1403;		
RX	MEDLINE=21235186; PubMed=11337471;		
RA	Bojocin A., Wincker P., Manger S., Jallion O., Malame K.,		
RA	Weissenbach J., Ehrlich S.D., Sorokin A.:		
RT	"The complete genome sequence of the lactic acid bacterium Lactococcus		
RT	lactis ssp. lactis IL1403."		
TL	Genome Res. 11:731-753(2001).		

```

CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE006396; AAK05758.1; -.
CC PIR; D86832; D86832.
CC HAMAP; MF_00551; -; 1.
CC DR InterPro; IPRO06082; PRK.
CC DR InterPro; IPRO06083; PRK URK.
CC DR InterPro; IPRO00764; Uridine_kin.
CC DR Pfam; PF00485; PRK; 1.
CC DR PRINTS; PR00478; PHRI1BKINASE.
CC DR PRINTS; PR00988; URIDINKINASE.
CC DR TrEMBL; TIGR00235; udk; 1.
CC KW Transferrase; Kinase; ATP-binding; Complete proteome.
CC FT NP_BIND 11
CC FT ATP (POTENTIAL).
CC SQ SEQUENCE 206 AA; 23740 MW; 82E38F052037DBEE CRC64;
-----
Alignment Scores:
Pred. No.: 1-56-22 Length: 206
Score: 454.00 Matches: 92
Percent Similarity: 62.50% Conservative: 43
Best Local Similarity: 42.55% Mismatches: 65
Query Match: 15.30% Indels: 16
DB: 1 Gaps: 5
-----
US-09-896-522-1 (1-1624) x URK_LACLA (1-206)
QY 154 CAGCGGCGCCCTTCGTATAGGGGTGAGCGGCGGACCTGCCAGCGGAGATGCAACGTGTGT 213
DB 2 LysLysThrLeuIleIleIleGlyValThrGlyLysSerAlaSerClyLysThrSerValSer 21
QY 214 GAGAAATCATGAGAGTCTCTGGGACAGAACGAGGTGAAACAGCGGACCGGAAAGTGTTC 273
DB 22 HisAlaIleLeuGluThrPhe---SerAsnGlu-----ArgIleAla 34
QY 274 ATCTCGAGCCAGGACAGGTTCTACAG-----GTCCGTGACGCGACAGAGAAAGGCC 324
DB 35 MetIleGluHisAspSerTyrTyrLysAspGlnSerHisSleThrPheIleGluLysThr 54
QY 325 AAGGCTTGAAGAAGCAGTACATTTTACCATTCAGCATGACCTTGTATGATGATTGGATG 384
DB 55 LysThr-----AsnTyrAspHisProLeuAlaPheAspThrAspTyrLeu 69
QY 385 CACAGGACTGTGAAGACATCGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 444
DB 70 IleAlaGlnLeuLysGluLeuGlnTyrGlyLysArgAlaValAspIleProIleTyrAspTyr 89
QY 445 GTGACACACTCAAGGTTACACAGAGCCACAGGAGGTCTACCCCTGGCGAGCTGTCTGTT 504
DB 90 AlaLysHisThrAspGlySerGlnGluThrTyrArgGlnGluProValAspValLeuIleVal 109
QY 505 GAGGCGACTTGTGGTGTCTTACAGCGCAGGAGATCCGAGCATGTTCCACTGTGGCGCTTC 564
DB 110 GluGlyIleLeuValLeuGlnLysAspGluLysArgLeuAlaGlyPheLeuMetAspIleLysIlePhe 129
QY 565 GTGACACCGCATCCGAGCTCAGGCTGTCTCGAAGAGTTCCTCGGAGACGTG---CGCCGA 621
DB 130 ValAspThrAspAspAspValArgIleIleArgAlaGlyIleArgArgAspIleGluGluLysG 149
QY 622 GGGAGGAGCCTGGAGCAGATTCTGACGCAAGTACACACACCTTGTGAAGCGGCGCTTCGAG 681
DB 150 GlyArgThrLeuAspSerValIleThrGlnTyrLeuAspAlaValAlaLysProMetTyrHis 169

```

```

OY 6632 GAGTTCCTGCTCCGCAAGAAGATGTGCGCAATGATCATCCGACGAGAGTGGAAT 741
Db 170 GlnpelleiegluProThryrlysrgrYrYrAlaaspVallllelProctluelYalSerasn 189
OY 742 ATGGTTCGCATCAACCTGATGTCGACGACATCCAGACATTTCTGAAT 789
Db 190 ThrValGlyValaspIleleThrrThryrIleValSerlleuasn 205

RESULT 12
URK_STAM STANDARD; PRT; 207 AA.
ID_URK_STAM
AC 0997N8;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Uridine kinase (EC 2.7.1.48) (Uridine monophosphokinase) (Cytidine
DE monophosphokinase).
GN UDK OR SAV1611 OR SAL439 OR MW1561.
OS Staphylococcus aureus (strain Mu50 / ATCC 700699),
OS Staphylococcus aureus (strain N315), and
OS Staphylococcus aureus (strain MM2).
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=158878, 158879, 196620;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mu50 / ATCC 700699, and N315;
RX MEDLINE=21311952; PubMed=11418146;
RA Kuroda M., Ohta T., Uchiyama I., Baba T., Yuzawa H., Kobayashi I.,
RA Cui L., Oguchi A., Aoki K.-I., Nagai Y., Lian J.-Q., Ito T.,
RA Kanamori M., Matsumaru H., Matuyama A., Murakami H., Hosoyama A.,
RA Mizutani-Uji Y., Takahashi N.K., Sawano T., Inoue R.-I., Kaito C.,
RA Sekimizu K., Hirkawa H., Kuhara S., Goto S., Yabuuchi J.,
RA Kanehisa M., Yamashita A., Oshima K., Funaya K., Yoshino C., Shiba T.,
RA Hattori M., Ogasawara N., Hayashi H., Hiramatsu K.;
RT "Whole genome sequencing of methicillin-resistant Staphylococcus
RT aureus."
RL Lancet 357:1225-1240(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MM2;
RX MEDLINE=22040717; PubMed=12044378;
RA Baba T., Takenuchi F., Kuroda M., Yuzawa H., Aoki K.-I., Oguchi A.,
RA Nagai Y., Iwana N., Asano K., Naito T., Kuroda H., Cui L.,
RA Hiramatsu K.;
RT "Genome and virulence determinants of high virulence community-
RT acquired MRSA."
RL Lancet 359:1819-1827(2002).
CC -1- CATALYTIC ACTIVITY: ATP + uridine = ADP + UMP.
CC -1- CATALYTIC ACTIVITY: ATP + cytidine = ADP + CMP.
CC -1- PATHWAY: Pyrimidine salvage pathway.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE URIDINE KINASE FAMILY.
CC
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboratio
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on it
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercia
CC entities requires a license agreement (See http://www.isb-sib.ch/announce
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AP003362; BAB57773.1; -
CC EMBL; AP003134; BAB42703.1; -
CC EMBL; AP004827; BAB95426.1; -
CC PIR; B89943; B89943.
CC HAMAP; MF_00551; -; 1.
CC InterPro; IPR006082; PRK.
CC InterPro; IPR006083; PRK URK.
CC InterPro; IPR00764; Uridine_kin.
CC Pfam; PF00485; PRK; 1.
CC PRINTS; PR00478; PHRIBLKINASE.
CC PRINTS; PR00988; URIDINKINASE.

```

[illegible]


```

QY      129  CGGCGCGGGA-----GGCGGACGTCGCGAC-----153
DB      40  LysThrGlyLysLysIleTyrThrLysGlyAspPro-ProTyrTyrAspLysGlyLys
QY      154  -----CAGCGGCGCTTCCTGATAGGGGTGAGCGCGGACCTGCGGAGGAGTGCAC 206
DB      59  sSerLeuLysHisProPheValIleGlyValCysGlyGlySerHisSerGlyLysThrTh
QY      207  CGTGTGTGAAGAATCATCGAGTGTCTGGGACAGACAGAGTGGAAACAGCGGACGGGAA 266
DB      79  rValAlaGlyLysIleValGluArgLeuGly-----11eProt 92
QY      267  GGTGTGATCTCTGAGCGGACAGAGTCTTCAAGGTCTTCAAGGTCTTCAAGGAGGAGGAGG 326
DB      92  pValThrIleLeuSerMetAspSerPheTyrLysValLeuThrProGluGluIleLysAl 112
QY      327  GGCCTTGAAGAAGACAGTACAAATTTTGACATCCAGATGCTTTGATATGATTTGATGCA 386
DB      112  aAlaHisGlySerArgTyrAsnPheAspGlyProAlaHisPheAspPheAspLeuLeuTy
QY      387  CAGGACTCTGAAGAAKATCGTGGAGGGGCAAAACGGTGGAGTGGCCGACCTATGATTTGT 446
DB      132  rGluValLeuLysArgLeuArgGluGlyLysSerValAspValProValTyrAspPheAs 152
QY      447  GACACACTGAAGGTTCACAGAGACCAAGGTGTCTACAGGTCTTCAAGGTGTCTTGTGA 506
DB      152  nThrHisSerArgAspProAsnSerLysMetMetTyrGlyAlaAspValLeuIlePheG 172
QY      507  GGGGATCTTGATGTTCTTACAGCCAGAGATCCGGGACATGTTCCAGCTGGCCTTTGCT 566
DB      172  uGlyIleLeuAlaPheHisAspGlyArgIleLysAsnLeuMetAspMetLysValPheVa 192
QY      567  GGACACCGGACTCCGACGCTGTCTCGAAGGTTCTCCGGGACCTGGCC---CGAGG 623
DB      192  lAspThrAspGlyAspLeuArgLeuAlaArgArgIleValAlaAspValThrAspArgG 212
QY      624  GAGGACCTGAGAGATTCTGACGACATACACACCTGTGTAAGCGGCGCTTCGAGGA 683
DB      212  yArgAspIleAspGlyIleMetGluGlnTyrPheThrPheValLysProAlaPheAspLy 232
QY      684  GTTCTGCTGCTCCGCAAAAGAGTATCCGATGTGATGATCCGACGAGAGTGGACAATAT 743
DB      232  sTyrIleAlaProCysMetAspSerAlaAspLeuIleValProArgGlyGlyGluAsnAs 252
QY      744  GGTGGCATCAACCTGATCTGTGACGACATC 774
DB      252  pValAlaIleAspMetIleValGlnAsnVal 262

RESULT 2
T21110
hypothetical protein F19B6.1b - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C/Accession: T21110
R/Thomas, K.
submitted to the EMBL Data Library, February 1996
A/Reference number: Z19375
A/Accession: T21110
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-555 <WIL>
A/Cross-references: EMBL:Z69635; NID:G1200023; PIDN:CAA93462.1; GSPDB:GN00022; CESP:F19B
A/Experimental source: clone F19B6
C/Genetics:
A/Gene: CESP:F19B6.1b
A/Map position: 4
A/Introns: 35/1; 83/3; 201/3; 267/3; 395/2; 511/3

Alignment Scores:
Pred. No.: 2.84e-32 Length: 555
Score: 561.50 Matches: 125
Percent Similarity: 63.10% Conservative: 46
Best Local Similarity: 46.13% Mismatches: 62

```

```

Query Match: 18.92% Indels: 38
DB: 2 Gaps: 6
US-09-896-522-1 (1-1624) x T21110 (1-555)

QY      12  TCCGACCTGGGCGCTGGG---CGGCGCGCGGCGGCGGAGGAGCGGCGGAGGAG 68
DB      54  SerProArgAlaIleGlyCysArgThrArgAspThrMetSerGlyGlyArgAlaGlu 73
QY      69  CCGATCGCGGAGCGGAGCGGAGCGGAGTGGCTTGGCGGAGCGGAGGAGGAGGAGGAG 128
DB      74  -----HisIleLeuLeuThrThr 79
QY      129  CGGCGCGGGA-----GGCGGACGTCGCGAC-----153
DB      80  LysThrGlyLysLysIleTyrThrLysGlyAspPro-ProTyrTyrAspLysGlyLys 99
QY      154  -----CAGCGGCGCTTCCTGATAGGGGTGAGCGGCGGACCTGCGGAGGAGTGCAC 206
DB      99  sSerLeuLysHisProPheValIleGlyValCysGlyGlySerHisSerGlyLysThrTh 119
QY      207  CGTGTGTGAAGAATCATCGAGTGTCTGGGACAGACAGAGTGGAAACAGCGGACGGGAA 266
DB      119  rValAlaGlyLysIleValGluArgLeuGly-----11eProt 132
QY      267  GGTGTGATCTCTGAGCGGACAGAGTCTTCAAGGTCTTCAAGGTCTTCAAGGAGGAGGAGG 326
DB      132  pValThrIleLeuSerMetAspSerPheTyrLysValLeuThrProGluGluIleLysAl 152
QY      327  GGCCTTGAAGAAGACAGTACAAATTTTGACATCCAGATGCTTTGATATGATTTGATGCA 386
DB      152  nThrHisSerArgAspProAsnSerLysMetMetTyrGlyAlaAspValLeuIlePheG 172
QY      387  CAGGACTCTGAAGAAKATCGTGGAGGGGCAAAACGGTGGAGTGGCCGACCTATGATTTGT 446
DB      172  rGluValLeuLysArgLeuArgGluGlyLysSerValAspValProValTyrAspPheAs 192
QY      447  GACACACTGAAGGTTCACAGAGACCAAGGTGTCTACAGGTCTTCAAGGTGTCTTGTGA 506
DB      192  nThrHisSerArgAspProAsnSerLysMetMetTyrGlyAlaAspValLeuIlePheG 212
QY      507  GGGGATCTTGATGTTCTTACAGCCAGAGATCCGGGACATGTTCCAGCTGGCCTTTGCT 566
DB      212  uGlyIleLeuAlaPheHisAspGlyArgIleLysAsnLeuMetAspMetLysValPheVa 232
QY      567  GGACACCGGACTCCGACGCTGTCTCGAAGGTTCTCCGGGACCTGGCC---CGAGG 623
DB      232  lAspThrAspGlyAspLeuArgLeuAlaArgArgIleValAlaAspValThrAspArgG 252
QY      624  GAGGACCTGAGAGATTCTGACGACATACACACCTGTGTAAGCGGCGCTTCGAGGA 683
DB      252  yArgAspIleAspGlyIleMetGluGlnTyrPheThrPheValLysProAlaPheAspLy 272
QY      684  GTTCTGCTGCTCCGCAAAAGAGTATCCGATGTGATGATCCGACGAGAGTGGACAATAT 743
DB      272  sTyrIleAlaProCysMetAspSerAlaAspLeuIleValProArgGlyGlyGluAsnAs 292
QY      744  GGTGGCATCAACCTGATCTGTGACGACATC 774
DB      292  pValAlaIleAspMetIleValGlnAsnVal 302

RESULT 3
T18629
hypothetical protein B0001.4 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 21-Jan-2000
C/Accession: T18629
R/Time, M.
submitted to the EMBL Data Library, February 1996
A/Reference number: Z18999
A/Accession: T18629
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA

```


A/Accession: T41020
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-454 <NR>
A/Cross-references: EMBL:AL023860; PIDN:CA19591.1; GSPDB:GN0068; SPDB:SPC162.11c
A/Experimental source: strain 972h-; cosmid c162
C/Genetics:
A/Status: SPDB:SPC162.11c
A/Map position: 3

Alignment Scores:
Pred. No.: 5,99e-25 Length: 454
Score: 459.00 Matches: 92
Percent Similarity: 62.33% Conservative: 47
Best Local Similarity: 41.26% Mismatches: 76
Query Match: 15.47% Indels: 8
DB: 2 Gaps: 2

US-09-896-522-1 (1-1624) x T41020 (1-454)

QY 124 AGCCCCCGCCGAGGCCGACCGTCCGACCGCGCCCTTCTGATAGGGTGAGCGGC 183
Db 9 SerSerAnpProThrYrGluProProTPrArgYValArgPheIleGlyIleAlaGly 28
184 GGCACTCCACCGGAGAGTGCACCGTGTGTGAAGAATCATGTGAGTTGCTGGACAGAAC 243
Db 29 ProSerIySerGlyYrThrSerValAlaGlnLeuIleValIySalAlaLeuAnLeuPro 48
QY 244 GAGGTGACACGCGGACCGGAGAGTGTGTCATCTGAGCCAGACAGATTCTACAGTGC 303
Db 49 Hs-----ValValIleLeuSerLeuAnSerPheYrYlySer 61
QY 304 CTGACGCGACAGGAGGAGGCCAAGCCTTGAAAGACATTCATTTTGACATCCAGAT 363
Db 62 LeuAnmIaGluGlnIySlyArgAlaPheAnmAnpYrYrAepPheAnpSerProGlu 81
QY 364 GCGTTGATATGATTTGATGCACAGAGCTCTGAAGAATCCGTGAGGCGAAGACGCTG 423
Db 82 AlAlaIleSprThrPheLeuPheValIyLeuLeuGlnLeuYsgInGlyArgIyVal 101
QY 424 GAGGTGCGGACCTTGTGATTTGTGACACACTCAAGGTTACAGAGACCGGTGTTCAC 483
Db 102 AspIleProIleYrSerPheAnmGlnHsAnmArgLeuProGlnThrAnThrLeuPhe 121
QY 484 CTTGGGACGCTGTCTGTGTGAGGACATCTGTGTTCACAGCCAGAGATCCGGAC 543
Db 122 GlyAlaSerIleIleIleLeuGlnGlyIlePheAlaLeuYrAepGluYrIleArgSer 141
QY 544 ATGTTCCACCTCGCCTCTTGTGTGACACCGACTCCGACGCTGTCTCGAAGATT 603
Db 142 LeuLeuAnpValSerValPheLeuAnpThrAspSerAspValCysLeuSerArgArgLeu 161
QY 604 CTCGGGACGTCGC---CGAGGAGGAGCCTGGAGAGATTCTGACGACAGTACCACC 660
Db 162 AsnArgAspIleAnYrArgGlyArgAspIleValGlyValLeuGlnGlnYrAnmIyS 181
QY 661 TTGCTGAGCGCGGCTTCGAGGACTTCTGCTGCCGCAAGAAAGTATGCCAGATGATC 720
Db 182 PheValIySprProSerYrGluAnPheValArgArgIleuSerYrThrAspLeuIle 201
QY 721 ATCCACGAGGAGGACAAATATGATGTGCATCACTGATCGTGACGACATCCAGAC 780
Db 202 ValProArgGlyArgAspAnmIyLeuAlaIleAspMetValIleAnPheIleArgArg 221
QY 781 ATTCTGAAT 789
Db 222 ThrLeuSer 224

RESULT 6
D86832
uridine kinase (EC 2.7.1.48) [imported] - Lactococcus lactis subsp. lactis (strain IL140
C/Species: Lactococcus lactis subsp. lactis
C/Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 03-Aug-2001

C/Accession: D86832
R/Bolotin, A.; Winkler, P.; Manger, S.; Jalllon, O.; Malarme, K.; Weissenbach, J.; Ehrlich
Genome Res. 11, 731-753, 2001
A/Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ssi
A/Reference number: A86625; WUID:21235186; PMID:1137471
A/Accession: D86832
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-206 <STO>
A/Cross-references: GB:AE005176; PID:g12724672; PIDN:AAK05758.1; GSPDB:GN00146
A/Experimental source: strain IL1403
C/Genetics:
A/Status: udk
C/Superfamily: uridine kinase
C/Keywords: phosphotransferase

Alignment Scores:
Pred. No.: 1.22e-24 Length: 206
Score: 454.00 Matches: 92
Percent Similarity: 62.50% Conservative: 43
Best Local Similarity: 42.59% Mismatches: 65
Query Match: 15.30% Indels: 16
DB: 2 Gaps: 5

US-09-896-522-1 (1-1624) x D86832 (1-206)

QY 154 GAGGCGCCTTCTGATAGGGTGAGCGGCGGACCTGCCAGGGAATGCACCGTGTGT 213
Db 2 LylYrThrLeuIleIleGlyValThrGlyYrSerAlaSerGlyYrThrSerValSer 21
QY 214 GAGAAATGATGAGTGTGCTGGGACAGACGAGTGAACAGCGGACGCGAGGTGTC 273
Db 22 HsAlAlaIleLeuGlnThrHe-----SerAnGln-----ArgIleAla 34
QY 274 ATCTGAGCCGACGACAGTTCACAG-----GTCTGACGCGACAGACAAAGCC 324
Db 35 MetIleGlnHsIerPheYrYrYrIySArgGlnSerHsIeLeuThrPheGlnGlnArgThr 54
QY 325 AAGGCTTGAAAGACAGTAAATTTTGACATCCGATGAGCTTGATATGATTGATG 384
Db 55 LylYrThr-----AnYrYrAspHsPProLeuAlaPheAspThrAspYrLeu 69
QY 385 CACAGAGCTGTGAAGAATCTGTGAGGCGAAGACCGGTGAGGCTGCAATGATTTT 444
Db 70 IleAlaGlnLeuIySerIleuGlnYrGlyArgAlaValAspIleProIleYrAspYr 89
QY 445 GTGACACACTCAAGTTACCAAGACACGAGTGTCTACCTCGCGACGTGTTCTGTT 504
Db 90 AlAluYrHsIthrYrSerGlnGlnIyThrYrArgGlnIyProValAspValIleuIleVal 109
QY 505 GAGGGATCTTGTTGTTCACAGCCGAGATCCGGAGATGTTCCACTGGCGCTTTC 564
Db 110 GlnGlyIleLeuValIleuGlnAspGlnArgLeuArgAspLeuSerAspIleYrIlePhe 129
QY 565 GTGACACCGACTCCGACCTCAGGCTGTCTGAAAGTCTCCGCGGACGTG---CGCGCA 621
Db 130 ValAspThrAspAspValArgIleIleArgArgIleArgArgAspIleGlnGlnArg 149
QY 622 GGGAGGACCTTGAGCAGATTCTGACGACGTCACCACTTGTTGTAAGCGGCTTCGAG 681
Db 150 GlyArgThrLeuAnpSerValIleThrGlnYrLeuAnpAlaValIySprOmetYrHs 169
QY 682 GAGTTCTGCTCGCGCAAGAAAGTATGCCAGTGTATATCCGACGAGGTGAGCAAT 741
Db 170 GlnPheIleGluProThrIySArgYrAlaAspValIleIleProGlnGlyValSerAn 189
QY 742 ATGTTGCATCACTGATCGTGACGACATCCAGACATTTCTGAAT 789
Db 190 ThrValGlyValAspIleIleThrThrIySIIeAlaSerIleLeuAn 205

RESULT 7
B89943
uridine kinase [imported] - Staphylococcus aureus (strain N315)

[illegible][illegible]

Qy 745 GTTGCCATCAACCTGATCTGTGACACATCCAGACATTCTG 786
Db 192 ValAlaIleAspLeuLeuThrThrLysIleAlaLysIleLeu 205

RESULT 9

A99008

uridine kinase (EC 2.7.1.48) [imported] - Streptococcus pneumoniae (strain R6)
C/Species: Streptococcus pneumoniae
C/Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 02-Nov-2001
C/Accession: A99008

R/Hokins, J.A.; Albom Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
e, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A/Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A/Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A/Reference number: A97872; MUID:21429245; PMID:11544234
A/Accession: A99008

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-212 <KOR>

A/Cross-references: GB:AE07317; PIDN:AAK9893.1; PID:G15458714; GSPDB:GN00174

C/Genetics:

A:Gene: udk

C:Superfamily: uridine kinase

C:Keywords: phosphotransferase

Alignment Scores:

Pred. No.:	1-71e-23	Length:	212
Score:	438.00	Matches:	88
Percent Similarity:	60.75%	Conservative:	42
Best Local Similarity:	41.12%	Mismatches:	68
Query Match:	14.76%	Indels:	16
DB:	2	Gaps:	4

US-09-896-522-1 (1-1624) x A99008 (1-212)

Qy 157 CGGCGCTTCTGATAGGGGTGAGCGCGGCGACCTCCAGCGGAAGTCGACCGTGTGAG 216
Db 4 ArgProIleIleIleGlyValThrGlyGlySerGlyGlyLysThrSerValSerArg 23

Qy 217 AAGTCAATGAGTTCCTGCGGACAGAGAGTGGAAACGCGGCGGAGAGTGTATC 276
Db 24 AlaIleuSer-----HisPheProAspGluLysIleSerMet 36

Qy 277 CTGAGCGGACGAGCTTCTACAG-----GTCTGACGCGGACGAGAGCGCAAG 327
Db 37 IleGluHisAspSerTyrrTyrrLysAspGlnSerHisLeuThrPheGluGluArgValLys 56

Qy 328 GCCTTGAAGAGACATCAATTGACCATCCAGATGCTTGTATGATTTGATGATC 387
Db 57 Thr-----AsnTyrrAspHisPheProAlaPheAspThrAspLeuMet 71

Qy 388 AGGACTGTGAAGACATCTGTGAGGCAAAACGTTGAGAGCGGACCTATGATTTGTG 447
Db 72 GluGlnIleLysGluLeuLeuAlaGlyLysProValAspIleProThrTyrrAspTyrrThr 91

Qy 448 ACACACTCAAGTTACAGACAGACACGAGTGTCTACCTCGGCGGAGTGTCTTTGAG 507
Db 92 GluHisThrArgSerSerLysThrTyrrArgGlnLysProGlnAspValIleIleValGlu 111

Qy 508 GGCATCTTGTGTCTTCTACAGCGAGAGATCCGGACATGTTCCACTGCGCTTGTG 567
Db 112 GlyIleLeuValIleGluLysAspLysArgLeuAspGlnMetAspIleLysIlePheVal 131

Qy 568 GACACGACCTCCGACGTCGAGCTGTCTGAAGATTTCTCGGACGCGG---CGCGGAGG 624
Db 132 AspThrAspAspAspValArgIleIleArgArgIleLysArgAspMetGluGluArgGly 151

Qy 625 AGGAGCTGAGAGATTTGTGACACACACCTTGTGAAGCGCGGCTTGAGAG 684
Db 152 ArgSerLeuAspSerValIleAsnGlnTyrrLeuGluValValLysProMetTyrrHisGln 171

Qy 685 TTTCGCTCGGACAAAGAGATATCGGATGTATCCACAGAGAGTGAATATG 744
Db 172 PheIleGluSerThrLysArgTyrrAlaAspIleValIleProGluGluValSerAsnThr 191

RESULT 10

S29374

uridine kinase (EC 2.7.1.48) - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein N2050; protein YNR012w
C/Species: Saccharomyces cerevisiae
C/Date: 28-May-1993 #sequence_revision 28-May-1993 #text_change 23-Mar-2001
C/Accession: S29374; S45134; S48349; S63338; S21361
R/Kern, L.

A/Title: The YRK1 gene of *Saccharomyces cerevisiae* encoding uridine kinase.
A/Reference number: S29373; MUID:90384830; PMID:2169608
A/Accession: S29374
A/Molecule type: DNA
A/Residues: 1-501 <KOR>
A/Cross-references: EMBL:X53998; NID:G4771; PIDN:CAA37946.1; PID:G4773
R/Vernhselt, P.; Aert, R.; Voet, M.; Voelckaert, G.
submitted to the EMBL Data Library, January 1994
A/Description: Twelve open reading frames revealed on the 23.6 kbp segment flanking the
A/Reference number: S45119
A/Accession: S45134

A/Molecule type: DNA

A/Residues: 1-501 <VER>

A/Cross-references: EMBL:X77395; NID:G496717; PID:G496728

A/Verhselt, P.; Aert, R.; Voet, M.; Voelckaert, G.

Yeast 10, 1355-1361, 1994

A/Title: Twelve open reading frames revealed in the 23.6 kb segment flanking the centromere

A/Reference number: S48338; MUID:95208356; PMID:7900425

A/Accession: S48349

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-501 <VE2>

A/Cross-references: EMBL:X77395; NID:G496717; PIDN:CAA54580.1; PID:G496728

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, January 1994

R/Aert, R.; Verhselt, P.; Voet, M.; Voelckaert, G.

submitted to the Protein Sequence Database, April 1996

A/Reference number: S62910

A/Accession: S63338

A/Molecule type: DNA

A/Residues: 1-501 <AER>

A/Cross-references: EMBL:Z71627; NID:G1302489; PID:E239785; PID:G1302490; MIPS:YNR012w

A/Experimental source: strain S288C

C/Genetics:

A:Gene: SGD:URK1

A/Cross-references: SGD:S0005295; MIPS:YNR012w

A/Map position: 14R

C/Keywords: phosphotransferase; transmembrane protein

F/154-170/Domain: transmembrane #status predicted <TM1>

F/419-435/Domain: transmembrane #status predicted <TM2>

Alignment Scores:

Pred. No.:	2.1e-23	Length:	501
Score:	437.50 <td>Matches:</td> <td>97</td>	Matches:	97
Percent Similarity:	53.76% <td>Conservative:</td> <td>46</td>	Conservative:	46
Best Local Similarity:	36.47% <td>Mismatches:</td> <td>88</td>	Mismatches:	88
Query Match:	14.75% <td>Indels:</td> <td>35</td>	Indels:	35
DB:	2	Gaps:	5

US-09-896-522-1 (1-1624) x S29374 (1-501)

Qy 85 GAGCGGAGATGCTTCCGCGGAGGCGAAGCTCGGAGCGCCGCGGAGCGGAC 144
Db 28 LysAlaAsnAlaValMetAspGlyGluValAspValLysLysThrLysLysSer 47

Qy 145 CGT-----CCGACACAGCGGCGCTTCTGATATGGGGTGAAGCGGCGACCTGCCAC 195
Db 145 CGT-----CCGACACAGCGGCGCTTCTGATATGGGGTGAAGCGGCGACCTGCCAC 195

```

Db      48 ArgTyrIleProProtrpThrThrProTyrIleIleGlyIleGlyAlaSerGlySer 67
QY      196 GGAAGATCCAGCCCTGTCTGTGAAGATCATATGAGTTGCTGGACAGAACGAGGTGGAACAG 255
      68 GilylserThserValAlaAlaIleValSerSerIle-----80
QY      256 CGGACGCGAAGGTGGTCATCTGAGCCAGCAGAGTTCTTACAAAGGTCCTGACGGCAGAG 315
      81 AsnValProTrpThrValIleuIleSerLeuAspSerPheTyrAsnProLeuGlyProGlu 100
QY      316 CAGAGCGCCAGGCGCTTGAAGAAGACAGTACATTTTTCAGCATCCAGATGCCCTTGATAT 375
      101 AspArgAlaArgAlaPheLysAsnGluTyrAspPheAspGluProAsnAlaIleAsnLeu 120
QY      376 GATTGATGACGAGACTCTGAAGAACATCTGGAGGGCAAAACGGTGGAGGCCGAC 435
      121 AspleuAlaTyrLysCysIleuAsnLeuLysGluIleLysArgThrAsnIleProVal 140
QY      436 TATGATTTTGTGACACACTCAGGTTACAGAG---ACCACGGTGGTCACTCCGCGGAC 492
      141 TyrSerPheValHisHisAsnArgValProAspLysAsnIleValIleTyrGlyAlaSer 160
QY      493 GTGGTTCTGTGAGGCGCATCTTGTTCTTACAGCCAGAGATCCGAGCATGTTCCAC 552
      161 ValValValIleGluGlyIleTyrAlaLeuTyrAspArgArgLeuLeuAspLeuMetAsp 180
QY      553 CTGCGCCTCTTCTGTGACACCGACTCCGAGTCCAGGCTGTCTGGAAGATTCCTCCGGAC 612
      181 LeuLysIleTyrValAspAlaAspLeuAspValCysLeuAlaArgAlaGlyLeuSerAla 200
QY      613 ---GTGCGCGGAGGAGGAGCTGAGCAGATTTCTGACAGTACACACCTTCTGTAAG 669
      201 IleValSerArgGlyArgAspLeuAspGlyCysIleGlnIleProGluLysPheValLys 220
QY      670 CCGGACCTTCGAGAGTTCCTGCTCCGACAAAGAGATGCGCATGTGATCATCCACGA 729
      221 ProAsnAlaValLysPheValLysProThrMetLysAsnAlaAspAlaIleIleProSer 240
QY      730 GAGGTGACATATGTTGCCATCAACCTGATCGTCGACGACATC-----774
      241 MetSerAspAsnAlaThrAlaValAsnLeuIleIleAsnHisIleLysSerLysLeuGlu 260
QY      775 -----CGGAG 780
      261 LeuLysSerAsnGluHisLeuArgGluLeuIleLysLeuGlySerSerProSerGlnAsp 280
QY      781 ATTCTGATGTGATCATC 798
      281 ValLeuAsnArgAsnIle 286

```

RESULT 11

G69728
uridine kinase udk - Bacillus subtilis
C/Species: Bacillus subtilis
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
C/Accession: G69728
R/Kunze, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter, C.; Bron, S.; Brouillet, S.; Bunsch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emerson, P.T.; Eutlian, K.D.; Errington, J.; Fabret, C.; Ferrari, E. Nature 390, 249-256, 1997
A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaux, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Luthers, Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelli, Rieger, M.; Rivolta, C.; Roche, B.; Roche, M.; Sadale, Y.; Sato, T.; Scanlon, A.; Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Serot, A.; Winters, P.; Wipet, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A.; Authors: Yoshikawa, H.F.; Zunsstein, E.; Yoshikawa, H.; Danchin, A.
A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: G69728

A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-211 <KUN>
A/Cross-references: GB:299117, GB:AL009126; NID:g2634966; PIDN:CAB14675.1; PID:g2635179
A/Experimental source: strain 168
C/Genetics:
A/Genes: udk
C/Superfamily: uridine kinase

Alignment Scores:

Pred. No.:	Length:	Matches:
6,39e-23	211	84
Score:	430.00	Conservative: 52
Percent Similarity:	63.26%	Mismatches: 63
Best local Similarity:	39.07%	Indels: 16
Query Match:	14.49%	Gaps: 4

US-09-896-522-1 (1-1624) x G69728 (1-211)

```

QY      154 CAGCGGCGCTTCTGTATAGGGGTAGCGGCGGACGCGACGCGAGCGGAAGTCGCTGT 213
      3 LysAsnProValValIleGlyIleAlaGlySerGlySerGlySerGlySerValThr 22
QY      214 GAGAGATCATGAGGTGGTGGGACAGAAAGAGTGAACAGCGGACGCGAGCGGAGTGT 273
      23 ArgSerIleTyrGlu-----GlnPheLysGlyHisSerIleLeu 35
QY      274 ATCTTGAGCCAGGACAGGTTCATCAG-----GTCTGACGCGACAGCAAGGCC 324
      36 MetIleGlnGlnAspLeuTyrTyrLysAspGlnSerHisIleuProPheGluGluAspLeu 55
QY      325 AAGCGCTTGAAGAGACAGTAAATTTTGAACATTCAGATGCCATGATATGATATGATG 384
      56 AsnThr-----AsnTyrAspHisProLeuAlaPheAspAsnTyrLeu 70
QY      385 CACAGGACTCTGAAGAACATCGTGAGGGCAAAACGTTGAGGTGCCGATGATTTT 444
      71 IleGluHisIleGlnAspLeuAsnTyrArgProIleGluLysProIleTyrAspTyr 90
QY      445 GTGACACACTCAAGGTTACAGACAGACCGAGTGTCTACCTCGGACGTTGTTT 504
      91 LysLeuHisThrArgSerGluGluThrValHisValGluProLysAspValIleIleLeu 110
QY      505 GAGGCGATCTGGTGTCTTACAGCCAGAGATCCGGAGATGTCACCTCGGCGCTTTC 564
      111 GlnGlyIleLeuValLeuGluAspLysArgLeuAspAspLeuMetAspIleLysLeuTyr 130
QY      565 GTGACACCGACTCGAGCTCAGGCTGTCTCGAAGAGTTCTCCGGAGCTG---CGCCGA 621
      131 ValAspThrAspAlaAspLeuArgIleIleArgArgIleMetArgAspIleAsnGluArg 150
QY      622 GGGAGGAGCTTGAGAGATTTCTGACGAGTACACACCTTCTGTAAGCCGCGCTTCGAG 661
      151 GlyArgSerIleAspSerValIleGluGlnTyrValSerValValArgProMetHisAsn 170
QY      682 GAGTTTCTGCTGCCGCAAGAAAGATATGCCAGTATCATCCACGAGAGAGTGACAT 741
      171 GlnPheValGluProThrLysArgTyrAlaAspIleIleIleProGluGlyGlnAsn 190
QY      742 ATGGTGCATCAACTGATCGTGCAGCAGCATCCATCCAGCATGTTCTG 786
      191 HisValAlaIleAspLeuMetValThrLysIleGlnThrIleLeu 205

```

RESULT 12

uridine kinase [imported] - Halobacterium sp. NRC-1
C/Species: Halobacterium sp. NRC-1
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 17-May-2002
C/Accession: F84277
R/Ng, W.V.; Kennedy, S.P.; Mahaitas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Laeky, S.; Leitbauer, B.; Koller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor, Jung, K.H.; Alam, M.; Freitas, T.
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A:Title: Genome sequence of *Halobacterium* species NRC-1.
A:Reference number: A84160; MUID:20504483; PMID:11016950
A:Accession: F84277
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-248 <STO>
A:Cross-references: GB:AE004437; NID:g10580750; PIDN:AA619586.1; GSPDB:GN00138
C:Genetics:
A:Gene: urk
C:Superfamily: uridine kinase

Alignment Scores:
Pred. No.: 5.58e-22 Length: 248
Score: 417.00 Matches: 83
Percent Similarity: 63.64% Conservative: 50
Best Local Similarity: 39.71% Mismatches: 58
Query Match: 14.05% Indels: 18
DB: 2 Gaps: 5

US-09-896-522-1 (1-1624) x F84277 (1-248)

```
OY      163  TTCTGATAGGGGAGCGGCGGCACTGCCAGCGGAAAGTCAGCTGTGTGAGAATGC 222
      29  PheAlaIleGlyIleAlaGlyIleThcIyAlaGlyIleThrValAlaAlaArgGluIle 48
OY      223  ATGAGTGTCTGGGACAGAAACAGAGTGGACAGCGGCGGCGGAAAGTGTGATCCTGAGC 282
      49  ThrAspAsnValGlyGluSer-----AlaThrIleuIlePro 60
OY      283  CAGGACAGGTTCTCAAGAGTCTTGACG-----GCAGACGAAAGGCCAGGCTTG 333
      61  LeuAspAsnTyTrpGluAspLeuSerAspArgProPheGluGluAlaAsnAla--- 79
OY      334  AAAAGACAGTACATTTTGAACATGCAGATGCCCTTGAATAGTATTGATGACAGAGACT 393
      80  -----AsnTyTrpAspHisProSerAlaPheGluTrpGluLeuLeuAlaGlyThrHis 95
OY      394  CTGAAGAATCTCTGGAGGCAAAACGATGAGGTGGCCGACCTATGATTTTGTGACACAC 453
      96  IleAspAlaLeuLeuSerGlyGlnSerIleGluMetProGlnTyTrpAspPheGluArgHis 115
OY      454  TCAGAGTTACAGACACCGGTGGTCTACCTTCGCGAGCTGTGTTGAGGGGATC 513
      116  ValArgIysAlaAspArgValAlaValGluProThrAspValIleValIleGluGlyIle 135
OY      514  TTGGTGTTCACACGAGACATCCGGGACATGTTCCACCGCGCCCTTCCTGAGACACC 573
      136  LeuAlaLeuSerAspGluThrValAlaAspMetLeuAspLeuHisIleTyTrpAlaGluThr 155
OY      574  GACTCCGACGTCAGCTGTCTCGAAGAGTTCTCCGGAGC---GTGCCCGCAGAGGAGGAC 630
      156  AspAlaAspValAlaGlyIleLeuArgArgIleGluArgAspValAlaGluArgGlyArgGlu 175
OY      631  CTGGAGCAGATTCTGACGCGAGTACACACCTTCGTGAAGCGCGCTTCGAGAGTTTGC 690
      176  LeuGluIlyAlaMetAspArgIleTyLeuSerThrValIysProMetHisGluGlnPheIle 195
OY      691  CTGGCCGACAAAGATGTCGCGATGATATCCACAGAGAGTGGACATTAAGTGGTCC 750
      196  GluProThrIysArgIleHisAlaAspIleIleProGluGlyAla---AsnSerValAla 214
OY      751  ATCAACCTGATCTGTCGACACATCCAG 777
      215  ValAsnLeuLeuGluGluIysValGln 223
```

RESULT 13

D86582
uridine kinase [imported] - *Chlamydomonas reinhardtii* (strain J138)
C:Species: *Chlamydomonas reinhardtii*, *Chlamydomonas reinhardtii*
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: D86582
R/Shitai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ig

Nucleic Acids Res. 28, 2311-2314, 2000
A:Title: Comparison of whole genome sequences of *Chlamydia pneumoniae* J138.
A:Reference number: A86491; MUID:20330349; PMID:10871362
A:Accession: D86582
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-222 <STO>
A:Cross-references: GB:BA000008; NID:g8979107; PIDN:BA98942.1; GSPDB:GN00142
A:Experimental source: strain J138
C:Genetics:
A:Gene: CP10735
C:Superfamily: uridine kinase

Alignment Scores:
Pred. No.: 1.25e-21 Length: 222
Score: 412.00 Matches: 84
Percent Similarity: 63.01% Conservative: 54
Best Local Similarity: 38.36% Mismatches: 59
Query Match: 13.89% Indels: 22
DB: 2 Gaps: 6

US-09-896-522-1 (1-1624) x D86582 (1-222)

```
OY      166  CTGATAGGGGTGACCGCGGCACTGCCAGCGGAAAGTCAGCTGTGTGAGAATGCATG 225
      8  IleIleGlyIleThrGlySerGlyAlaGlyIleThrLeuThrGlnAsnIleIys 27
OY      226  GAGTGTCTGGGACAGAAACAGAGTGGAAACGCGGACCGGAAAGTGTGATCCTGACCCG 285
      28  GluIlePheGlyIleAsp-----ValSerValIleCysGln 39
OY      286  GACAGGTTCTACAG-----GTCTGACGCGCAGACAGAGGCGCAAGGCTTGAA 336
      40  AspAsnTyTrpIysAspArgSerHisTyTrpProGluGluArgAlaAsnLeuIle---- 58
OY      337  GGAAGTACATTTTGAACATCCAGATGCCCTTGAATAGTATTGATGACAGAGACTTG 396
      59  -----TrpAspHisProAspAlaPheAspAsnAspLeuIleSerAspIle 74
OY      397  AAGAATGCTGGAGGCGCAAAACGATGAGGTGGCCGACCTATGATTTTGTGACACATCA 456
      75  LysArgLeuIysAsnAsnGluIleValGlnAlaProValPheAspPheValIleGluIys 94
OY      457  AGG---TTACAGAGACCAACGCGGTGTCTACCTTCGCGAGAGTGGTGTGAGGCGATC 513
      95  ArgSerIysThrGluIleGluThrIleTyTrpProSerIleValIleLeuValGluGlyIle 114
OY      514  TTGGTGTTCACACGAGACATCCGGGACATGTTCCACCTGCGCTTCTGAGACACC 573
      115  LeuValPheGluAsnGlnGluLeuArgAspLeuMetAspIleArgIlePheValAlaPthr 134
OY      574  GACTCCGACGTCAGCTGTCTCGAAGAGTTCTCCGGAGCTGCC---CGAGGAGGAGAC 630
      135  AspAlaAspGluArgIleLeuArgArgMetValArgAspValGlnIleGluIysAspSer 154
OY      631  CTGGACAGATTCTGACGAGTACACACCTTCGTGAAGCGCGCTTCGAGAGTTTCTGC 690
      155  ValAspCysIleMetSerArgTyTrpLeuSerMetValIysProMetHisGluIysPheIle 174
OY      691  CTGGCCGACAAAGATGTCGCGATGATATCCACAGAGAGTGGACATTAAGTGGTCC 750
      175  GluProThrArgIysArgIleHisAlaAspIleIleValHisGlyAsnTyArgGluIleAsnVal 194
OY      751  ATCAACCTGATCTGTCGAC-----CACATCCAGACATTTGAAATGATGAC 795
      195  ThrAsnIleLeuSerGlnIysIleLysAsnHisLeuGluAsnAlaLeuIleuIleuIleuIle 213
```

RESULT 14

E72041
uridine kinase CP0011 [imported] - *Chlamydomonas reinhardtii* (strains CW029 and AR39)
C:Species: *Chlamydomonas reinhardtii*, *Chlamydomonas reinhardtii*
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-May-2000
C:Accession: E72041; E81622

R:Kaltman, S., Mitchell, W., Marxthe, R., Lammel, C., Fan, J., Ollinger, L., Grimwood, J.
Nature Genet. 21, 385-389, 1999
A:Title: Comparative genomes of *Chlamydia pneumoniae* and *C. trachomatis*.
A:Reference number: A72000; MUID:99206606; PMID:10192388
A:Accession: E72041
A:Molecule type: DNA
A:Residues: 1-222 <ARR>
A:Cross-references: GB:AE001655, GB:AE001363, NID:g4377039, PIDD:AA08874.1; PID:g4377039
A:Experimental source: strain CWL029
R:Read, T.D.; Brumham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey,
C.; Dodson, R.; Gwinn, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
Nucleic Acids Res. 28, 1397-1405, 2000
A:Title: Genome sequences of *Chlamydia trachomatis* MoPn and *Chlamydia pneumoniae* AR39.
A:Reference number: AB1500; MUID:20150255; PMID:10684935
A:Accession: B81622
A:Molecule type: DNA
A:Residues: 1-222 <REA>
A:Cross-references: GB:AE002165, GB:AE002161; NID:g7188948, PIDD:AAF37907.1; PID:g7188948
A:Experimental source: strain AR39, HL cells
A:Genetics:
A:Gene: CP0735, CP0011
A:Superfamily: uridine kinase

Alignment Scores:	
Pred. No.:	1,256-21
Score:	412.00
Percent Similarity:	63.0%
Best Local Similarity:	38.36%
Query Match:	13.89%
DB:	2
Length:	222
Matches:	84
Conservative:	54
Mismatches:	59
Indels:	22
Gaps:	6

Oy		CTGATAGGGGTGAACGCCGCCTACTCCAGCGGGAAAGTCGACC GTGTGTGAGAAGATCATG	225
dY	166	CTGATAGGGGTGAACGCCGCCTACTCCAGCGGGAAAGTCGACC GTGTGTGAGAAGATCATG	225
bD	8	IleIleGIlylErnmGlLySerGIylAclGLyLThrTmrLeEThrcInnaEnIlEyS ::: ::: ::: :::	27

Oy	226	GAGTTGCTGGACACGAA	CGAGGTGGAA	CAGCGG	CAGGAAGGTGTCTCCTGAGCCAG	285
Ob	28	GuilePhecllyGluasp	-----	-----	-ValserVallicCyasin	39

	QY	286	GACAGGTTCTAACAAG-----GTCCCTGA	CGGCAGAGCAGAAGGCCAACGCTTGAAA	336
			:::: :: :: :: :: :: :: :: :: :: ::	:: :: :: :: :: :: :: :: :: ::	
	b	40	AspdadenthyrlylvspasparSerHtistyrThrpproglucgualalasnleule---	58	
			::: :: :: :: :: :: :: :: :: :: ::	:::	

Oy 337 GGACAGTACATTGTGACCATCCAGATGCCCTTGTAATGATTGTATGCACAGACTCTG 396

db 59 -----TTPASPhISPcRAsPaIapneAsnAmSpleuleuLeleSeArAile 74
::: :::
::: :::

OY 397 AAGAA CATCGTGAGGGCAAAA CGGTGAGGTGCACCTATGATTTTTGTGACA CACTCA 456
:: :: :: :: ::
DB 75 LYSATCLeuLYSAsnGlnIurIValIGlnAlProValPheAsnPheValIleuGIValsn 94

[illegible][illegible][illegible][illegible]

691 CTGCCGCAAGAAGTATGCCGATGTGTCATCCCAAGAGAGTGACATATGTTGCC 750

QY	751	ATCAACCGTATCGACGAG-----CACATCCAGGACATTTGCAATGCTGAC	795
Db	195	ThrsnllleueuerglnlysllelylsasnhllsneuglasmnlaleuGluserap	213

RESULT 15
G70101
uridine kinase (udk) homolog - Lyme disease spirochete

C.Prepared: 13-Feb-1998 #sequence_revision 13-Feb-1998 #extl_change 17-Nov-2000
 C.Date: 13-Feb-1998 #sequence_revision 13-Feb-1998 #extl_change 17-Nov-2000
 C.Accession: G0101, 146979
 R.Frazer, C.M., Casjens, S., Huang, W.M., Sutton, G.G., Clayton, R., Lathigra, R., White

A: Authors: Smith, H.O.; Venter, J.C.
N: Nature 390, 580-586, 1997
C: Cullen, R.E.; Gairland, S.; Fujii, C.; Colton, M.D.; Horst, K.; Roberts, K.; Hatch, B.
P: PNAS 94, 1033-1038, 1997
R: Rasmussen, J.; Gairland, S.; Horst, K.; Roberts, K.; Hatch, B.
S: Science 275, 122-124, 1997
T: Tishler, P.V.; Gairland, S.; Horst, K.; Roberts, K.; Hatch, B.
U: U.S. Environmental Protection Agency, 1997
V: Venter, J.C.; Smith, H.O.; Gairland, S.; Horst, K.; Roberts, K.; Hatch, B.

Accession	Gene	Protein	Accession	Gene	Protein
A:Accession: G70101	A:Reference number: 970100; MIMD:98065943; PMID:9403865	A:Status: preliminary; nucleic acid sequence not shown; translation not shown	A:Accession: G70101	A:Reference number: 970100; MIMD:98065943; PMID:9403865	A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A: Molecule type: DNA
A: Residues: 1-207 <KLE>
A: Cross-references: GB:AE00115; GB:AE00783; NID:G2687879; PIDN:AAC66392.1; PID:G2687880
A: Experimental source: strain B31

A. Reference number: 224339. MUID: 97372541. PMID: 9228761

A.Accession: 146979
A.Status: preliminary; translated from GB/EMBL/DBJ
A.Molecule type: DNA
A.Residues: 2-207 <BOU>

A:Cross-references: EMBL:X57449, NID:g1405440; PIDN:CAA6081.1; PID:g1405442
A:Experimental source: strain HB19, ssp. sensu stricto
C:Genetics:
A:Gene: uidK

C:Superfamily: uridine kinase	
Alignment Scores:	
Read No.: 1	59e-21
Length:	207

Score:	410.50	Matches:	79
Percent Similarity:	59.81%	Conservative:	49
Best Local Similarity:	36.92%	Mismatches:	59
Query March	13.84%	Indels:	27

DB:	2	Gaps:	4
US-09-896-522-1 (1-1624) x G70101 (1-207)			

QY CTGATAGGGGATGACGGCGGCACTGCCAGCGGGAAGTGCACCCGTGTGGAAGATCATG 225
166 :::::|||||
Db 5 TlelleglyTleersglyLyseryserolystrmThrvaiValserylsleSer 24
TlelleglyTleersglyLyseryserolystrmThrvaiValserylsleSer 24

QY 226 GAGTTGCTGGGAGACGAACGAGGTGCAACACGGCAGCGGAGGATGTCATCCGTGAGCAG 285
 25 Gluphelliisprogu-----PheValLeuIleSerGln 35

QY 286 GACAGTTCTCAAGGCTTCTGCAGCAGACGAAGAAGCCATTAAGGACAATAAC 345
||| :||:
Db 36 AspaentYtyrlyrsSerVal-----GIyaepTyr 45

QY	346	-----AATTTGACCATCCAGAGCCCTTTGATTAATGATTGGATG	384
Db	46	GIuHISGIuPHeSerLySvaIAsnPhEaSpHISProASpAlaPhEaSPaSnInLeuPHe	65

QY 385 CACAGGACTCTGTAAGAACATCTGTGGAGGGCAAAACGGTGGAGGTGCCGACCTATGATTTT 444

Db 66 TGTGTAHLSLEULYSASNLLEULYLYSASNSERPROILEASPETPROLEUYNIRAPHE 85

QY 445 GTGACACTCAAGGTTACCAAGAGACCAAGGTGTCTACCCCTGGGACGTGGTCTGTTT 504
:::|||||
Db 86 ILeSnHslyGsrGgInleuNtrvalleuValProHrrProValIleVal 105
:::|||||

OY 505 GAGGCACTTGTGTTCTACAGCCAGGAGATCCGGGACATGTTCCACTTGCGCCTTC 564
|||||||:::||||:::|||||:::|||||:::|||||

[illegible]

```
Db      106 GluGlyIleMetIlePheValGluGluArgValArgAsnLeuIleAspLeuIleTyr 125
QY      565 GTGGACACCGACTCCGACTCGCTCTCGAAGATTCTCCGGAGAGTG--CGCCGA 621
      ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      126 IleAspThrProAsnAspIleArgPheIleArgIleuArgArgAspIleSerIleArg 145
QY      622 GGGAGGAGCTGAGACAGATTCTGACGAGTACACACCTTCGTGAAGCCGGCCTTCGAG 681
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      146 GlyArgThrValGluSerValIleAspGlnTyrLeuAsnThrThrArgTrpGlyTyrTyr 165
QY      682 GAGTTCTGCCCTCCGACAAAGATATGCCGATGTGATCATCCACGAGAGAGTGGACAT 741
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      166 ArgPheIleGluProThrLysGlnTyrIleAspLeuIleIleProGluGlyGlyHisAsn 185
QY      742 ATGGTTGCCATCAACTGATCGTCGACGACATCCAGACATT 783
      |||||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      186 AspLysAlaLeuTyrValLeuSerThrPheLeuIysSerLeu 199
```

Search completed: November 25, 2003, 07:40:36
Job time : 48.9992 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus.n2p model1

Run on: November 25, 2003, 07:30:54 ; Search time 153.943 Seconds
(without alignments)
3891.516 Million cell updates/sec

Title: US-09-896-522-1

Perfect score: 2967
Sequence: 1 gfgggggtcgccctccgagctc.....ccaggtctgtgfggggagcag 1624

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 673684 seqs, 184443283 residues

Total number of hits satisfying chosen parameters: 1347368

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODE=frame+n2p.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US09896522/runat_21112003_184107_2917/app_query.fasta_1.2830
-DB=Published Applications_AA -QFMT=fasten -SUFFIX=n2p.rapb -MIMATCH=0.1
-LOOPEXT=0 -LOOPEXT=0 -UNITS=bites -START=1 -END=1 -MATRIX=blosum62
-TRANS-human0.cgi -LIST=45 -DOCCALIGN=200 -THR SCORE=ext -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEARSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09896522_@CGN_1_1_129_@runat_21112003_184107_2917
-NCPU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOC=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Published Applications_AA:*
1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/1/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/1/pubppa/US09A_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubppa/US09B_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/prodata/1/pubppa/US10A_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubppa/US10B_PUBCOMB.pep:*
15: /cgn2_6/prodata/1/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep:*
18: /cgn2_6/prodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
------------	-------	-------------	--------	----	-------------

1	1450	48.9	277	9	US-09-896-522-2
2	925	31.2	337	10	US-09-925-300-1160
3	428.5	14.4	125	9	US-09-896-522-5
4	327	11.0	60	12	US-10-029-386-28765
5	320.5	10.8	124	9	US-09-896-522-6
6	276	9.3	231	9	US-09-896-522-4
7	192.5	6.5	797	15	US-10-156-761-10907
8	177	6.0	595	10	US-09-738-973-187
9	177	6.0	595	10	US-09-854-133-187
10	177	6.0	595	10	US-10-144-649A-187
11	172.5	5.9	503	14	US-10-078-547-2
12	169.5	5.8	265	11	US-09-975-719-369
13	169	5.7	1207	15	US-10-024-368-5
14	164	5.6	737	12	US-10-078-547-24
15	164	5.6	737	12	US-10-017-161-1350
16	163	5.5	1535	15	US-10-189-971-14
17	163	5.5	1570	15	US-10-189-971-12
18	162	5.5	1182	15	US-10-024-368-6
19	159.5	5.4	1063	12	US-10-017-161-1624
20	158.5	5.4	375	15	US-10-156-761-14622
21	158.5	5.4	1593	15	US-10-189-971-4
22	158.5	5.4	1628	15	US-10-189-971-2
23	156.5	5.3	877	12	US-10-237-496-90
24	156.5	5.3	877	12	US-10-242-074-90
25	156.5	5.3	877	12	US-10-242-505-90
26	156.5	5.3	877	12	US-10-242-574-90
27	156.5	5.3	877	12	US-10-243-261-90
28	156.5	5.3	877	12	US-10-243-282-90
29	156.5	5.3	877	12	US-10-243-402-90
30	156.5	5.3	877	12	US-10-243-431-90
31	156.5	5.3	877	12	US-10-244-164-90
32	156.5	5.3	877	12	US-10-244-972-90
33	156.5	5.3	877	12	US-10-197-942-90
34	156.5	5.3	877	12	US-10-238-166-90
35	156.5	5.3	877	12	US-10-238-193-90
36	156.5	5.3	877	14	US-10-041-770-2
37	156.5	5.3	877	15	US-10-245-103-90
38	156.5	5.3	877	15	US-10-245-107-90
39	156.5	5.3	877	15	US-10-245-113-90
40	156.5	5.3	877	15	US-10-245-771-90
41	156.5	5.3	877	15	US-10-245-851-90
42	156.5	5.3	877	15	US-10-245-883-90
43	156.5	5.3	877	15	US-10-237-535-90
44	156.5	5.3	877	15	US-10-238-183-90
45	156.5	5.3	877	15	US-10-238-283-90

ALIGNMENTS

RESULT 1
US-09-896-522-2
; Sequence 2, Application US/09896522
; Patent No. US20020055161A1
; GENERAL INFORMATION:
; APPLICANT: Glucksmann, Maria A.
; TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
; FILE REFERENCE: 381552001700
; CURRENT APPLICATION NUMBER: US/09/896,522
; CURRENT FILING DATE: 2001-06-28
; PRIOR APPLICATION NUMBER: 60/216,503
; PRIOR FILING DATE: 2000-06-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-896-522-2

Alignment Scores: 4.26e-101 Length: 277
Pred. No.: 1450.00 Matches: 277
Score:

Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 48.87% Indels: 0
DB: 9 Gaps: 0
US-09-896-522-1 (1-1624) x US-09-896-522-2 (1-277)
QY 94 ATGGCTTCGGCGGAGGAGAGCTGGAGAGCCCGCGGAGCGACCGTCCGAC 153
DB 1 MetAlaSerAlaGlyGluAspGlySerProAlaProGluAlaAspArgProHis 20
QY 154 CAGCGCCCTTCCTGATAGGGGTAGCGCGGCACTCGCAGCGGGAAGTCAGCTGTGT 213
DB 21 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 40
QY 214 GAGAGATCATGAGTGTCTGGGACAGAACGAGGTGGAACAGCGGCAAGCGGAGTGTCT 273
DB 41 GluLysIleMetGluLeuGlyGlnAsnGluValGluGlnArgGlnArgLysValVal 60
QY 274 ATCTGAGCCGAGACAGGTTTACAAAGTCTTGAACGGACAGAGCGAAGGCGCTTG 333
DB 61 IleLeuSerGlnAspArgPheTyLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 80
QY 334 AAAGACAGTACAAATTTGACCATCCAGATGCTTTGATATGATTGATCAGAGACT 393
DB 81 LysGlyGlnTyrraAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
QY 394 CTGAGAAACATGTGGAGGGGCAAAACGGTGGAGGTGCCGACTTATGATTGTGACAC 453
DB 101 LeuLysAsnIleValGluGlyLysThrValGluValProThrTyrraAspPheValThrHis 120
QY 454 TCAGGTTCACGAGACCGGTGTCTACCCCTCGACCGGTCTTGTGAGGGATC 513
DB 121 SerArgLeuProGluThrThrValValTyrraAlaAspValValLeuPheGluGlyLe 140
QY 514 TTGGTGTCTCACGCGAGAGATCCGGGACATGTTCCACCTGCGCCCTCTTGTGACAC 573
DB 141 LeuValPheTyrrSerGlnGluIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
QY 574 GACTCCGACGTGAGCGTGTCTCGAAAGTTTCCGGGACGTGCCCGGAGGGAGACTG 633
DB 161 AspSerAspValArgLysSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
QY 634 GAGAGATTCGAGCGAGTACACACCTTGTGAAGCGGCGCTTGGAGAGTTGTGCTG 693
DB 181 GluGlnIleLeuThrGlnTyrrThrThrPheValLysProAlaPheGluGluPheCysLeu 200
QY 694 CCGACAAAGATATGCGATGTATCATCCACGAGAGGTGACAAATATGTTGCATC 753
DB 201 ProThrLysLysTyrrAlaAspValIleIleProArgGlyValAlaAsnMetValAlaIle 220
QY 754 AACCTGATGTCAGACATCCAGACACTTGTGAATGGTGCATCTGCAATGGCACCGA 813
DB 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlnAspIleCysLysTyrrHisArg 240
QY 814 GGAAGGTTCATATGGCGGAGGTACAGAGGACCTTTCTGAGCGAGGGAGCAACCTGGG 873
DB 241 GlyLysSerAsnGlyArgSerTyrrLysArgThrPheSerGluProGlyAspHisProGly 260
QY 874 ATGCTGACCTTGGCAACGCTGCATTTGGAGTCCAGACAGACACCCCA 924
DB 261 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerArgProHis 277
RESULT 2
US-09-925-300-1160
; Sequence 1160, Application US/09925300
; Patent No. US20020151681A1
; GENERAL INFORMATION:
; APPLICANT: Craig Rosen,
; APPLICANT: Steve Ruden
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA101
; CURRENT APPLICATION NUMBER: US/09/925,300

CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05988
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1890
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1160
; LENGTH: 337
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (38)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (46)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (155)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (169)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-925-300-1160
Alignment Scores:
Pred. No.: 2,02e-61 Length: 337
Score: 925.00 Matches: 200
Percent Similarity: 73.53% Conservative: 25
Best Local Similarity: 65.36% Mismatches: 58
Query Match: 31.18% Indels: 23
DB: 10 Gaps: 6
US-09-896-522-1 (1-1624) x US-09-925-300-1160 (1-337)
QY 12 TCCGACCTCGGCGCTGGCGGGCGCGCGCGGAGGAGGGCGCGGAGACCGG 71
DB 54 SerAspAlaGlyValIleGlyValAlaSerValAlaArgThrGlySerGly----- 70
QY 72 ATGCGCGGAGCGGAGCGGAGAGTGGCTTCGCGGAGGCGGAGACTGGAGACCCCGC 131
DB 71 ---ArgArgGlyGlyValAsnHisGlyArgGlyGlnArgAlaAspProAlaGluProPro 89
QY 132 GCCGAGGCGGACCGCTCCGACCGAGCGGCGCTTCTCGAATAGGGGTAGCGGCGGCTCG 191
DB 90 AlaAlaGlnArg-----ArgArgAlaLeuProTyrrArgArgHis-GlyGlyThrAl 106
QY 192 CAGCGGAGATCGACCGTGTGTGAGAAAGATCATGAGTGTCTGGGACAGAACGAGGTGGA 251
DB 106 AsnGlyLysSerSerValCysAlaLysIleValGlnLeuGlyGlnAsnGluValAs 126
QY 252 ACAGCGGACGCGAGAGTGTGTATCTGAGCGGACAGAGCAGGTTCTACAGTCTCGACGC 311
DB 126 pTyrrArgGlnLysGlnValValIleLeuSerGlnAspSerPheTyrrArgValLeuThrSe 146
QY 312 AGACGAGAGGCCAAGCGCTTGAAGAAGACATTTTACCATCCAGATGCCCTTTGA 371
DB 146 rGluGlnLysValAlaLysValLeuLys**GlnPheAsnPheAspHisProAspAlaPheAs 166
QY 372 TAATGATTTGATCAGACGCTGTGAAGAACATTCGTGAGGGGCAAAACGTGAGAGTGGC 431
DB 166 PheGlnLys***IleLeuLysThrLeuLysGluIleThrGluGlyLysThrValGlnIleThr 186
QY 432 GACCTATGATTTTGTGACACACTCAAGGTTTACAGAGACCAAGGTGTACCTGCGGA 491
DB 186 oValTyrraPhePheValSerHisSerArgLysGluGluThrValThrValTyrrProAlaAs 206
QY 492 CGTGTCTCTGTTTGAAGGCATCTTGCTGTCTTACAGCCGAGAGATCCGGACATGTTCCA 551
DB 206 pValValLeuPheGluGlyIleLeuValaPheTyrrSerGlnGlnValArgAspLeuPheG 226
QY 552 CCGGCGCTCTTGTGTGACACCGGACCTCGGACGTGTCTCGAAGGTTCTCGGGGA 611

[illegible]

```

Db      61 G1yG1yAsPAsnHisValAlaIleAsPLeuIleValGlnHisIleGlnSerIleLeuAsn 80
Qy      790 GGTGACATCTGCACAAATGGCAGCGAGGGGTCCAAATGGCGGAGCTCAACAGCGACCTTT 849
      :::
Db      81 GluGlyLeuSerSerGlnHisThrAsnTyrlrMeValAsnArgSerTyrlrAsnGlyThrPhe 100
Qy      850 TCTGAGCCAGGGGAGCCACCTCTGGAGTGTGACCTCTGGCAACCGTGCACATTTGGAGTCC 909
      |||||
Db      101 SerGluProGlyAspHisProGlyTyrlrProSerGlyAsnArgGlnHisIleGluSer 120
Qy      910 AGCAGCAGACCCAC 924
      |||||
Db      121 SerSerArgProHis 125

RESULT 4
US-10-029-386-28765
; Sequence 28765, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR GI
; FILE REFERENCE: AEOICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 28765
; LENGTH: 60
; TYPE: PRF
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO CHR9.1
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3
; OTHER INFORMATION: EXPRESSED IN HETAL, SIGNAL = 3.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 3.3
; OTHER INFORMATION: SWISSPROT HIT: P52623, EVALUE 8.00e-26
US-10-029-386-28765

Alignment Scores:
Pred. No.: 1,91e-16 Length: 60
Score: 327.00 Matches: 60
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 11.02% Indels: 0
DB: 12 Gaps: 0

US-09-896-522-1 (1-1624) x US-10-029-386-28765 (1-60)
Qy      745 GTTGCATCAACCTGATCTGTGAGACATCCAGACATTTCTGAATGTGACATCTGCAA 804
      |||||
Db      1 ValAlaIleAsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLys 20
Qy      805 TGGCACCAGAGGAGGTCTCAATGGGCGGAGCTCAACACGACCTTTCTGAGCCAGGGGAC 864
      |||||
Db      21 TrpHisArgGlyGlySerHisnGlyArgSerTyrlrLysArgThrPheSerGluProGlyAsp 40
Qy      865 CACCTGGAGATGCTGACCTCTTGCAACACGTCACATTTGGATGTCAGCAGACACCCAC 924
      |||||
Db      41 HisProGlyMetLeuThrSerGlyAsnArgSerHisIleGluSerSerSerArgProHis 60

RESULT 5
US-09-896-522-6
; Sequence 6, Application US/09896522
; Patent No. US20020055161A1

```

```

GENERAL INFORMATION:
APPLICANT: Glucksmann, Maria A
TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
FILE OF INVENTION: US98THEROF
FILE REFERENCE: 381552001700
CURRENT APPLICATION NUMBER: US-09/896,522
CURRENT FILING DATE: 2001-06-28
PRIORITY APPLICATION NUMBER: 60/216,503
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ. ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 124
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Consensus amino acid sequence
US-09-896-522-6

Alignment Scores:
Pred. No.: 7,12e-16 Length: 124
Score: 320.50 Matches: 62
Percent Similarity: 67.19% Conservative: 24
Best Local Similarity: 48.44% Mismatches: 37
Query Match: 10.80% Indels: 5
DB: Gaps: 2

US-09-896-522-1 (1-1624) x US-09-896-522-6 (1-124)

QY 166 CTGATAGGGGTGACCGCGGCCTGACCGCGGAATGTCACCGTGTCGAGAATCATG 225
Db 1 TlEllelylllealaglyserglyserglysthrinthlealaargylsileval 20
QY 226 GAGTTGCTGGAGACAAGACAGAGGTGAACACGCCGACCGGACCGGAAGTGTTCATCCTGACGCAG 285
Db 21 GLMtleuLsn-----lysProglnglnllyvalIlelleSerGln 36
QY 286 GACACGGTTTCAACAGGTCCTGACGGCACAGCCAAAGGCCAAGCCTTGAAGAAGACAGTAC 345
Db 37 AsPasnlyKtyrlyAsrpleuSerGluleuAerMetglugluArglysgluAsnaSnlyr 56
QY 346 AATTGTGACCATCCAGATGCGCTTGATATGATTGATGACACAGGACTGTGAAGACATC 405
Db 57 AsnPheasrhierprosrAlraRheasBrpheasrleuTyrglnHisleuLys**Leu 76
QY 406 GTGGAGGGCAAACGGGTGAAGGTGCCGACCTTAATGATTTTGTGACA--CACTCAAGGTTA 462
Db 77 LysasnDlyySseValglnuValProIleTyraErphelystHnHisIarGArglyS 96
QY 463 CCAAGACCAACGGGTGTAACCTCGGAGACGGTGTCTGGTAGGGGAGCATTTGGTGTTC 522
Db 97 AspdlunhValthrIleagluproAlasrvalIlelleuglulglylerTyralau 116
QY 523 TACAGCCAGAGAGATCCGGGACATG 546
Db 117 TyrAspplurJlgIleArGAsrpleu 124

RESULT 6
US-09-896-522-4
Sequence 4, Application US/09896522
Patent No. US20020055161A1
GENERAL INFORMATION:
APPLICANT: Glucksmann, Maria A.
TITLE OF INVENTION: 57658, A NOVEL HUMAN URIDINE KINASE AND
FILE OF INVENTION: US98THEROF
FILE REFERENCE: 381552001700
CURRENT APPLICATION NUMBER: US-09/896,522
CURRENT FILING DATE: 2001-06-28
PRIORITY APPLICATION NUMBER: 60/216,503
PRIOR FILING DATE: 2000-06-30
NUMBER OF SEQ. ID NOS: 6
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 4
```

```

; LENGTH: 231
; TYPE: PR
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Consensus amino acid sequence
US-09-896-522-4

Alignment Scores:
Pred. No.: 1, 92e-12 Length: 231
Score: 276.00 Matches: 68
Percent Similarity: 50.95% Conservative: 39
Best Local Similarity: 32.38% Mismatch: 81
Query Match: 9.30% Indels: 22
DB: 9 Gaps: 5

US-09-896-522-1 (1-1624) x US-09-896-522-4 (1-231)

QY 166 CTGATATAGGGTGAAGCGCGGCACCTGCCAGCGGGAAGTGCACCCGTGTGAGAAATCATG 225
      :::::::::::::::::::: ::::::::::::::::::::
DB 1 ValIleGlyValAlaIleAlaYserSerGlyAlaGlyIleThrThrValAlaArgArgIleVal 20

QY 226 GAGTGTGTGGACAGACGAGGTTG-----GAAACAGCGGCAG 261
      ::::::::::::::::::::
DB 21 SerIlePheGlyArgGluGlyValProAlaAlaGlyIleGluGlyAlaSerProAspSerAsn 40

QY 262 CGAAGAGTGTATCTCTGACGAGACAGACAGATTCTTCAAGAGTCTTGACGCGACAGACAGAG 321
      ::::::::::::::::::::
DB 41 ThrGlyAspSerPheLeuArgLeuAspArgPheTyrMetCAspLeuHisIleGluIAspArg 60

QY 322 GCCAAGGCTTGAAAGACAGACATCAATTTTGACCATTCAGATGCTTGATTAATGATTGG 381
      ::::::::::::::::::::
DB 61 LysArgAlaGlyAlaAsnLysHisIleTyrIlePhePheSerProGluAlaAsnAspPheAspLeu 80

QY 382 ATGCACAGCACTCTGAAAGAACATCGTGAGGCAAAACGATGAGAGTGCACGACCTATGAT 441
      ::::::::::::::::::::
DB 81 LeuTyrGluValPheLysGluLeuLysGluGlyLysSerValAspIleAspIleTyrAsn 100

QY 442 TTGTGTACACACTCAAGGTTACCAG-----ACCAGGTGTACCT 486
      ::::::::::::::::::::
DB 101 HisValThrGlyGlyLysArgProAspGlyGlnGluProGlyThrPheThrAspTyrPro 120

QY 487 -----GCGACGTGGTCTGTGTGAGGCACTTGATGTTTACAGCCAG 531
      ::::::::::::::::::::
DB 121 GluLeuIleGluGlyAlaAspValIleuValIleGluGlyLeuHisIleValTyrAspGlu 140

QY 532 -----GAGATCCGGGACATGTTCCACTGCGCCTCTTGATGACACCGACTCCGAC 582
      ::::::::::::::::::::
DB 141 ArgGluValAsnValIleGluLeuLeuAspLeuLysIleTyrValAspProAspIleAsp 160

QY 583 GTGAGGTGTCTGAGAAAGTCTCCGGGACGTG---CGCGAGGGAGGAGACCTGGAGACG 639
      ::::::::::::::::::::
DB 161 LeuGluLeuAlaArgLysIleGluIleArgAspMetAlaGluIleGlyHisIleSerIleGluGly 180

QY 640 ATTGTGACGACATCACCACTCTGTGAAAGCGGCTTCGAGAGATTCTGCTTCGCGACA 699
      ::::::::::::::::::::
DB 181 ValLeuAspSerIleGluLysArgArgLysProAspTyrValAsnTyrIleAlaIleProGln 200

QY 700 AAGAAGTATGCGATGATCATCCACAGA 729
      ::::::::::::::::::::
DB 201 PheSerTyrAlaAspLeuIleIleGlnArg 210

RESULT 7
US-10-156-761-10907
; Sequence 10907, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHITYUKI
; APPLICANT: HATTORI, MASAHIRA

```

```

; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 10907
; LENGTH: 797
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-10907

```

Alignment Scores:

```

Pred. No.: 5.34e-06 Length: 797
Score: 192.50 Matches: 115
Percent Similarity: 35.03% Conservative: 36
Best Local Similarity: 26.68% Mismatches: 158
Query Match: 6.53% Indels: 122
DB: 15 Gaps: 22

```

US-09-896-522-1 (1-1624) x US-10-156-761-10907 (1-797)

```

QY 1181 CTCAGTGCCTAGAGGAGCTTTTAACCGC-----AACGAGCCTAAG 1140
DB 333 lIethrAspneuHsAlaAlaLeuAsnLysAlaLysAlaAspIleMetAspGlyProAla 352
QY 1139 TGGCTGAAACCTCAGGAGAGCGCTGTGATGCCAGCAAGTTGAGTCTGAGTGACACAT 1080
DB 353 SerMetIethrArgLeuProAlaIleIleGln----- 364
QY 1079 CTGAGTTTCACTCTCAGTGAAGAGCGCTCGCTTAACACTCCCTGGGGTGGCGG 1020
DB 365 -----HsGln-----ArgArgAlaLysThrLeuAlaLeu-ThrGlnLeuSerAl 379
QY 1019 AGAGGAAGCAGTGGGTGGGTGGGCGGCCCGGCTCAGTCCAGTCCCTGAACACATGCCG 960
DB 379 aLeuHsAlaThrSerSerArgProAlaPro-----CysG1 391
QY 959 GCGGAGACCTGCGCTGAGGCTCGGAGCCCTCAGTGGGGTCTGTGACTGCCAA 900
DB 391 yArg-----SerAlaProSerProLeuProAlaSerGly-----AlaGlnAr 406
QY 899 TGTGACCTTTGCCAGAGGTCAGATCCAGAGGTGTCCTGGCTGAGAAAGTCCGC 840
DB 406 gSerValIleCysGlyAlaSerValAlaThrAlaArgThrGlyGlnLysGlyPheSe 426
QY 839 TTGTAGTCCGCCCATTTGACC-----CTCCTGGTGCATTTGCAG 798
DB 426 rValSerSerThrArgValThrLeuArgAgaGlnGlnArgLeuValTrpAlaLeuAlaG1 446
QY 797 ATGTCACTTCAAGATGTCTGATGTCTGCACGATGATGATGATGATGATGATGATGATG 728
DB 446 nArgSerHisValProCysAlaGlyCysPro----- 456
QY 737 TCACATCTCTGTTGGATGATCATCATCTTCTTTGTCCGACGAGCAACTCTCT-- 680
DB 457 -----GluSerArgLeuArgAlaLeuArgProLeuProSerProse 470
QY 679 -----CGAAGCCGGCTTCAAGAGGTGGTGTCTCGTCAGAAATC 639
DB 470 rLeuLeuProProAlaAlaThrArgProAlaSerProArg----- 483
QY 638 TGTCTCAGATCCCTCCGCGGACGCTCCCGAGAACTTTGAG-----ACAGCCTG 585
DB 484 -----ProAlaAlaAlaAlaAlaAlaArgProArgAlaLeuProSerProThrThrSerAl 502
QY 584 ACGTCCGAGTCCGCTCAGAGAGGCGCAGGTGGAACATGTCCTGGATCTCTGGCTG 525
DB 502 aAlaArgAlaThrSerProSerMet-----ThrLeuProProProValTrpAr 518

```

```

QY 524 TAGAACCAAGATGCTTCAACAGAACCAAGCTCCGACGGGTAGA----- 479
DB 518 gArgProAlaArgSerSerArgProProAlaArgSerSerArgArgThrAlaSerPr 538
QY 478 -----CCACCGGTCTCTGTGAACCTTGAGTGT 450
DB 538 oThrProThrArgCysSerValTrpSerSerSerProSerProAlaThrThrArgSerVa 558
QY 449 GTCACAAAATCATAGTCCGACACTCCACCGTTTCCCTCCACGATGTTCTTCAAGTC 350
DB 558 lSerAlaSer---ProThrProProProSerArgLysProProArgSerThr----- 574
QY 389 CTGTGCATCAATCATATTCAAAGCATCTGATGTGTCAAATTTACTGTCTTCAAG 330
DB 575 -----ArgArgProProSerAlaSerSerThrThrSerArgSerAr 588
QY 329 GCCTTGACCTTCTGTCTGCGCGTCAAGACCTTGAGAACCTGTCTGGCTCAGAGTACC 270
DB 588 gArgThrThrSerArgThrSerSerSerThrArgArgArg-----Pr 602
QY 269 ACCTTCGCGTCCGCTGTTCCACTCTGTCTGTCGCCAGCAACTCATGATCTTCTACAC 210
DB 602 oProThrSerProAlaSerProPro-----ProArgProProSerArgThrSerSe 619
QY 209 ACCGTGACCTCCCGCTGCGAGTGCAGCCGCTCACCCCTATCAGAAAGGCGCGTGTGC 150
DB 619 rAlaSerSerAlaAlaTrp---ThrSerArgSerThrSerSerArgArgAlaSerArgAr 638
QY 149 GAGCGGTGCGCTCCGCGCGGGGCTCTGCACTTTCGCTCCCGCAGAACCATCTCG 90
DB 638 gAlaSerArgThrProArgArgValSerArgSerArgSerArgSerThrPro-----Ar 655
QY 89 GCCTCCGCTCCGCGCATGGGTGTCCTCCGCGCGCCCTTCCCGGCGCCGCG---CGCGC 33
DB 655 gProPro---ProArgAlaGlySerPro-----AlaArgThrArgAl 668
QY 32 CCGCCAGCGCGGAGGTGCGAGGCGCACCCCA 2
DB 668 aArgProArgProArgAlaArgSerAlaPro 678

```

RESULT 8

```

US-09-738-973-187
; Sequence 187, Application US/09738973
; Patent No. US20020110563A1
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Fling, Steven P.
; APPLICANT: Mohamath, Radooh
; APPLICANT: Algate, Paul A.
; APPLICANT: Secrist, Heather
; APPLICANT: Indrias, Carol Yoseph
; APPLICANT: Benson, Darin R.
; APPLICANT: Elliott, Mark
; APPLICANT: Mannion, Jane
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C9
; CURRENT APPLICATION NUMBER: US/09/738,973
; CURRENT FILING DATE: 2000-12-14
; NUMBER OF SEQ ID NOS: 587
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 187
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-738-973-187

```

Alignment Scores:

```

Pred. No.: 7.34e-05 Length: 595
Score: 177.00 Matches: 110

```


Percent Similarity: 33.42%
 Best Local Similarity: 27.64%
 Query Match: 5.97%
 DB: 10
 Gaps: 22

US-09-896-522-1 (1-1624) x US-09-738-973-187 (1-595)

```

QY 5 GGTGCGCTCCGACCTCGGCGCTGGGCGGCGCGCGCGCGGGAAGGCGGCGCGG 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 228 GlyAlaSerAlaProGluSerGlnAlaGlyGly---GlyPro---ArgGlyArgAlaArg 245
QY 65 GGACCCGATGCGCGGAGCGGAGCGCGAGATGCTCGCGGAGCGGAGATGCGGAGA 124
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 GlyPro---ArgGlnGln-GlyArgArgArgHisGlyThrGlnArgAlaArgGlyPr 263
QY 125 GCGCGCGCGCGGGA-----GCGCGGACC 145
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 263 oProGlnAlaArgGluGluGlyProArgAlaThrThrIleuGlyLeuGlyThrPr 283
QY 146 GTCCGCA-----CCAGCG----- 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 283 oSerGlyGluGlnArgAlaSerGlnSerGlnAlaLeuProAlaLeuAlaGlyAlaAla 303
QY 159 -----GCCCTTCGATAGGGGTGAGCGCGCGCACTGCGCGGAAAGTCCA 205
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 aAlaHisAlaHisAlaIleProGlyAlaGly--ProAlaAlaAlaProValGlyArg 322
QY 206 CCGTGTGAGAAAGATCATGAGTTGCTGGAGACAAAGAGTGGAGCGGAGCGGGA 265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 323 Gly-----ArgArgGly-GlyTTPArgGlyGlyArgArgGlySerAlaGlyAlaG 340
QY 266 AGGTGTCATCTCGAGCGAGACAGGTTCTACAGGTCCTGACGACGAGAGAGGCA 325
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 340 YGlyGly-----GlyArgGlyGlyArgGlyGlyArgGlyArgGlyArg 351
QY 326 AGGCTTGAAGAAGACATGACATTTTGAACCATCCAGATCCTTGATATGATTTGATGC 385
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 351 gGly----- 352
QY 386 ACAGAGACTGGAAGAACATCGTGGAGGCAAAAGGTGAGGCGGACCTATGATTTTG 445
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 353 -----GlyGlyArgGlyGlyGlyAla 360
QY 446 TGACACACTCAAGTTACAGAGACACAGGTGATCTACCTCGAGACGTGTTCTGTTTG 505
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 -----GlyArgGlyGly----- 364
QY 506 AGGCACTTGTGTTCTACAGCCAGAGATCCGGACATGTTCCACTGCGCTCTTCG 565
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 365 -----GlyAlaAlaGlyProArgGlu---GlyAlaSerSerProGlyAlaArgAr 380
QY 566 TGAACACGACTCCGACGTCAGGCTGTCTCGAAGAGTTCTCCGGAGCGTGGCGGAGGA 625
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 380 gGlyGluGlnArgArgArg-----GlyArgGlyPro---Pr 391
QY 626 GGAAGCTGGAGCAGATTCTGACGAGTACACACACTCGTGAAGCGCGCTTCGA----- 680
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 391 oAlaAlaGlyAlaAlaGlnValSerAlaArgGlyArgArgAlaArgGlyGlnAlaG 411
QY 681 -----GGAGTTTGTCTGCGGACAAAGAGATAGCCGATGTGATCTCC 724
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 411 yGluGlnAlaGlnAspGlyLeuLeuProArgGlyArgAspArgLeu-----Pr 427
QY 725 CACGAGAGTGGACAAATGTTGTCATCAACCTGATGCGACACATCCAGGACATTC 784
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 427 oLeuArgProGlyAspAlaAsn-----GlnArgAlaGlnArgProGly----- 441
QY 785 TGAATGTGACATCTGCAATGAGCACCGAGAGGATCCATGAG-----CGAGCTACA 838
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 442 -----ProProArg-GlyGlyGlnHisGlyProValAspAlaSer 454
QY 839 AGCGGACCTTTTCTGAGCCAGGAGACACCTGGAGATGCTGACTCTGGCAACGAGTCA 898
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 454 eAlaProAspThrSerProProAlaHisProAlaArgTrpValSerGlnAlaArgGlnA 474
QY 899 ATTTGAGTGCACAGACAGACCCCACTGAGG---GCTGCCGAGCCTCAGGACGATCTC 955
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 474 rGluTTPArgGlnPheArgValGlyGlyGlyPheProProProProSerArgProP 494
QY 956 CCGCC-----CGCATGTGTTCAGGAGTGAAGCTTGGAGCGCC 997
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 494 roAlaValLeuLeuProLeuLeuArgLeuAlaCysAlaGly---AppProGlyAlaGThra 513
QY 998 ACCCACACCACTGCTTCCTTCGCGGACCCCAAGGAGAGTGT 1041
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 513 rProGlyProArgArgProAlaArgArgProArgGlyGluLeu 527

```

RESULT 9

```

US-09-854-133-187
; Sequence 187, Application US/09854133
; Publication No. US20020183499A1
; GENERAL INFORMATION:
; APPLICANT: Lodges, Michael J.
; APPLICANT: Mohamath, Raodoh
; APPLICANT: Henderson, Robert A.
; APPLICANT: Benson, Darin R.
; APPLICANT: Secrist, Heather
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C10
; CURRENT APPLICATION NUMBER: US/09/854,133
; CURRENT FILING DATE: 2001-05-11
; NUMBER OF SEQ ID NOS: 735
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 187
; LENGTH: 595
; TYPE: PRT
; ORGANISM: Homo sapien
US-09-854-133-187

```

Alignment Scores:

```

Pred. No.: 7,34e-05 Length: 595
Score: 177.00 Matches: 110
Percent Similarity: 33.42% Conservative: 23
Best Local Similarity: 27.64% Mismatches: 118
Query Match: 5.97% Indels: 147
DB: 10 Gaps: 22

```

US-09-896-522-1 (1-1624) x US-09-854-133-187 (1-595)

```

QY 5 GGTGCGCTCCGACCTCGGCGCTGGGCGGCGCGCGGGAAGGCGGCGCGG 64
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 228 GlyAlaSerAlaProGluSerGlnAlaGlyGly---GlyPro---ArgGlyArgAlaArg 245
QY 65 GGAAGCTGGAGCAGATTCTGACGAGTACACACACTCGTGAAGCGCGCTTCGA----- 124
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 246 GlyPro---ArgGlnGln-GlyArgArgArgHisGlyThrGlnArgArgGlyPr 263
QY 125 GCGCGCGCGCGGA-----GCGCGGACC 145
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 263 oProGlnAlaArgGluGluGlyProArgAlaThrThrIleuGlyLeuGlyThrPr 283
QY 146 GTCCGCA-----CCAGCG----- 158
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 283 oSerGlyGluGlnArgAlaSerGlnSerGlnAlaLeuProAlaLeuAlaGlyAlaAla 303
QY 159 -----GCCCTTCGATAGGGGTGAGCGCGGCACTGCGAGGGAAGTCCA 205
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 aAlaHisAlaHisAlaIleProGlyAlaGly--ProAlaAlaAlaProValGlyArg 322
QY 206 CCGTGTGAGAAAGATCATGAGTTGCTGGAGACAAAGAGTGGAGCGGAGCGGGA 265
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 323 Gly-----ArgArgGly-GlyTTPArgGlyGlyArgArgGlyGlySerAlaGlyAlaG 340
QY 266 AGGTGTCATCTCGAGCCAGAGACAGGTTCTACAGGTCCTGAGACGAGAGAGGCA 325
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db      340 yglYgly-----GlyArgGlyGlyArgGlyArgGlyArg 351
Qy      326 AGGCCTTGAAGAAGACAGTACATTTTGACCATCCAGATGCTTTGATTAATGATTTGATGC 385
Db      351 gglY----- 352
Qy      386 ACAGAGCTCTGAAGAATCGTGAGGGCAAAACGGTGAGTGCCGACCTATGATTTTG 445
Db      353 -----GlyGlyArgGlyGlyGlyGlyAla----- 360
Qy      446 TGACACACTCAAGTTACAGAGACCAAGGTGTTACCTCGCGAGCGTGTCTGTTTG 505
Db      361 -----GlyArgGlyGly----- 364
Qy      506 AGGGCATCTTGTTGTTCTACAGCCAGAGATCCGGGACATGTTCCACTTCGCGCTCTTCG 565
Db      365 -----GlyAlaAlaGlyProArgGlu--GlyAlaSerSerProGlyAlaArgAr 380
Qy      566 TGACACCCGACTCCGAGCTGCTCGAAGAGTTTCGGGAGACGTGCGCGAGGGA 625
Db      380 gglYglnGlnArgArgArg-----GlyArgGlyPro--Pr 391
Qy      626 GGGACCTGAGACAGATTCTGACGACGTACACACTTCGTGAAGCCGCGCTTCGA----- 680
Db      391 oAlaAlaGlyAlaAlaGlnValSerAlaArgGlyArgAlaArgGlyGlnArgAlaG1 411
Qy      681 -----GGAATTCTGCTGCGCGCAAAAGATAGCCGATGTATCATCC 724
Db      411 yglGlnAlaGlnAspGlyLeuLeuProArgGlyArgAspArgLeu-----Pr 427
Qy      725 CACGAGAGTGACCAATATAGTTGCCATCAACCTGATCGTGACAGACATCCAGGACATTC 784
Db      427 oLeuArgProGlyAlaSerAlaAsn-----GlnArgAlaGlnArgProGly----- 441
Qy      785 TGAATGTGACATCTGCAAAATGGACCGAGAGAGTTCATGGG-----CGAGCTACA 838
Db      442 -----ProProArg-GlyGlyHisGlyProValAsnAlaSerS 454
Qy      839 AGCGGACCTTTTGTGAGCCAGGGAGCAACCTGTGGATGCTGACCTCTGGCAAAACGGTCA 898
Db      454 eAlaProAspThrSerProArgHisAspProArgArgTrpValSerGlnGlnArgGlnA 474
Qy      899 ATTGGAGTCCAGACAGACCCCACTGAGAG--GCTGCCGAGCCCTCAGGACAGGCTC 955
Db      474 rglLeuTrpArgGlnPheArgValGlyGlyPheProProProProSerArgProp 494
Qy      956 CCGCC-----CGGATGTGTGTTTCAGGAGCTGAGGCTGGGGAGCGCC 997
Db      494 roAlaValLeuLeuProLeuLeuArgLeuAlaCysAlaGly---AspProGlyAlaThrA 513
Qy      998 ACCCACACCCACTGCTTCTCTCGCGCGCACCCCAAGGGGAGTGT 1041
Db      513 rGProGlyProArgArgProAlaArgArgProArgGlyGlnLeu 527

RESULT 10
US-10-144-649A-187
; Sequence 187, Application US/10144649A
; Publication No. US20030118599A1
; GENERAL INFORMATION:
; APPLICANT: Lodes, Michael J.
; APPLICANT: Wang, Tongtong
; APPLICANT: Pan, Liqun
; APPLICANT: Algate, Paul A.
; APPLICANT: McNeill, Patricia D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 210121.475C11
; CURRENT APPLICATION NUMBER: US/10/144,649A
; CURRENT FILING DATE: 2002-08-21
; NUMBER OF SEQ ID NOS: 749
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 187
; LENGTH: 595

```

```

; TYPE: PRT
; ORGANISM: Homo sapien
US-10-144-649A-187

Alignment Scores:
Pred. No.: 7,34e-05 Length: 595
Score: 177.00 Matches: 110
Percent Similarity: 33.42% Conservative: 23
Best Local Similarity: 27.64% Mismatches: 118
Query Match: 5.97% Indels: 147
DB: 15 Gaps: 22

US-09-896-522-1 (1-1624) x US-10-144-649A-187 (1-595)

Qy      5 GGTGCGCTCCGACCTTGCGCGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 64
Db      228 G1yAlaSerAlaProGlnSerGlnAlaGlyGly--GlyPro--ArgGlyArgAlaArg 245
Qy      65 GGAACCGGATGCGCGGAGCGGAGCGCGGAGATGAGCTTCGCGGAGGCGAAGACTCGGAGA 124
Db      246 GlyPro-----ArgGlnGln-GlyArgArgArgHisGlyThrGlnArgArgArgGlyPr 263
Qy      125 GCGCCGCGCGCGA----- 145
Db      263 oProGlnAlaArgGlnGlnGlyProArgAspAlaThrThrIleLeuGlyLeuGlyThrPr 283
Qy      146 GTCCGCA-----CGAGG----- 158
Db      283 oSerGlyGlnGlnArgAlaAspGlnSerGlnAlaLeuProAlaLeuAlaGlyAlaAla1 303
Qy      159 -----GCCCTTCTGATPAGGGGTGAGCGCGCGGCACTGCGCGGAGAGTCA 205
Db      303 aAlaHisAlaHisAlaIleProGlyAlaGly--ProAlaAlaAlaProValGlyGlyArg 322
Qy      206 CCGTGTGTGAAGAATCATGAGTGTGCTGGGACAGAACAGAGTGGAACAGCGGACCGGA 265
Db      323 Gly-----ArgArgGly-GlyTrpArgGlyGlyArgGlyGlySerAlaGlyAlaG1 340
Qy      266 AGGTGTCTCTCTGAGCCAGGACAGGTTCTACAAGTCTGACGCGGACGAGAGAGGCCA 325
Db      340 yglYgly-----GlyArgGlyGlyArgGlyArgGlyArg 351
Qy      326 AGGCCTTGAAGAAGACAGTACATTTTGACCATCCAGATGCTTTGATTAATGATTTGATGC 385
Db      351 gglY----- 352
Qy      386 ACAGAGCTCTGAAGAATCGTGAGGGCAAAACGGTGAGTGCCGACCTATGATTTTG 445
Db      353 -----GlyGlyArgGlyGlyGlyGlyAla----- 360
Qy      446 TGACACACTCAAGTTACAGAGACCAAGGTGTTACCTCGCGAGCGTGTCTGTTTG 505
Db      361 -----GlyArgGlyGly----- 364
Qy      506 AGGGCATCTTGTTGTTCTACAGCCAGAGATCCGGGACATGTTCCACTTCGCGCTCTTCG 565
Db      365 -----GlyAlaAlaGlyProArgGlu--GlyAlaSerSerProGlyAlaArgAr 380
Qy      566 TGACACCCGACTCCGAGCTGCTCGAAGAGTTTCGGGAGACGTGCGCGAGGGA 625
Db      380 gglYglnGlnArgArgArg-----GlyArgGlyPro--Pr 391
Qy      626 GGGACCTGAGACAGATTCTGACGACGTACACACTTCGTGAAGCCGCGCTTCGA----- 680
Db      391 oAlaAlaGlyAlaAlaGlnValSerAlaArgGlyArgArgAlaArgGlyGlnArgAlaG1 411
Qy      681 -----GGAATTCTGCTGCGCGCAAAAGATAGCCGATGTATCATCC 724
Db      411 yglGlnAlaGlnAspGlyLeuLeuProArgGlyArgAspArgLeu-----Pr 427
Qy      725 CACGAGAGTGACCAATATAGTTGCCATCAACCTGATCGTGACAGACATCCAGGACATTC 784
Db      427 oLeuArgProGlyAlaSerAlaAsn-----GlnArgAlaGlnArgProGly----- 441

```

QY 785 TGAATGATCATCTGCAAAATGACACCGAGAGGGTCCATATGCG-----CGAGCTACA 838
 Db 442 -----ProProArg-GlyGlyHisGlyProValAsnHisAsers 454
 QY 839 AGCGACCTTTTCTGAGCCAGGAGACACCTGGGATCTGACCTTGGCAACGGTCCAC 898
 Db 454 ealAProAspThrSerProProArgHisProArgArgTyrValSerInGlnArgLna 474
 QY 899 ATTGGAGTCCAGACGACCCCACTGAGG---GCTGCCAGCCTCAGGCGAGTCTC 955
 Db 474 rgleuTPArgGlnPheArgValGlyGlyPheProProProProSerArgProP 494
 QY 956 CCGCC-----CGCATGTGTCTTCAGGAGTGAAGCTGGAGCGCC 997
 Db 494 roAlaValLeuLeuProleuLeuArgLeuAlaCysAlaGly---AspProGlyAlaThrXa 513
 QY 998 ACCACACCCACTGCTTCTCTCGGCGACACCCAGGAGGTCTT 1041
 Db 513 rGProGlyProArgArgProAlaArgArgProArgGlyGluLeu 527

RESULT 11

US-10-078-547-2
 ; Sequence 2, Application US/10078547
 ; Publication No. US20020199211A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Narayanaswamy Ramesh
 ; APPLICANT: Miguel A. de la Fuente
 ; APPLICANT: Ines M. Anton
 ; APPLICANT: Raif S. Geha
 ; TITLE OF INVENTION: WIP, A WASP-Associated Protein
 ; FILE REFERENCE: 1242.1022-005
 ; CURRENT FILING DATE: 2002-02-19
 ; PRIOR APPLICATION NUMBER: US/10/078.547
 ; PRIOR FILING DATE: 1997-12-23
 ; PRIOR APPLICATION NUMBER: 60/066,533
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 503
 ; TYPE: PRT
 ; ORGANISM: Human
 US-10-078-547-2

Alignment Scores:

Pred. No.: 0.000154 Length: 503
 Score: 172.50 Matches: 99
 Percent Similarity: 34.51% Conservative: 18
 Best Local Similarity: 29.20% Mismatches: 156
 Query Match: 5.86% Indels: 66
 DB: 14 Gaps: 14

US-09-896-522-1 (1-1624) x US-10-078-547-2 (1-503)

QY 1036 TCCCTGGGGTGGCCCGAGAGAGCAAGTGGGTGGGCTCCCAAGGCTAGTCC 977
 Db 154 SerProGlyHisArg-----SerGlyProProGluProGlnArg 166
 QY 976 CTGAACACACATGCGCGGCGAGACCTGCCCTGAGGCTCGGACGCCCTCAGTGGGATC 917
 Db 167 ---Asn-ArgMetProProProAlaGProAlaGlySerIleProAspSerIleProPr 185
 QY 916 TGCTGCTGACTCCAAATGTAACCGTTTGCCAGAGTCAAGATCCAGGGTGTCCCTG 857
 Db 185 oProValProSerThrProArgProAlaGlnSerSerLeuHisAsnArgGlySerProPr 205
 QY 856 GCTCAGAAAAGTCCGCTTGAGCTCGGCCCATTTGACCCCTCTCGGTGCATTGTGAGA 797

Db 205 oValProGlyGlyProArgGlnProSerPro---GlyProThrProProPhe----- 222
 QY 796 TGTACACATTCAAAATGTCCTGATGTCGACACGATCAGTGTGATGGCAACCATATTGT 737
 Db 223 -----ProGlyAsnArg-Gly 227
 QY 736 CCACTCTCGTGGAGATGATCATCGCATACTCTTTGCG-----GCA 692
 Db 228 ThrAlaLeuGlyGlySerIleArgGlnSerProLeuSerSerSerProPheSer 247
 QY 691 GGCAGAACTCTCTGAAGCGCGCTTACAGAGGTGTGTACTCGTCAGAAATGCTCTCCA 632
 Db 248 AsnArgProProLeuProProThrProSerArgAlaLeuAspAspIlyProProPro 267
 QY 631 GGTCCCTCCCTCGGCGGACGTCCTCGGAGAACTTTGACAGACCCCTGAGCGAGTCCG 572
 Db 268 ProProProValGlyAsnArgPro-----SerIleHisArgGluAla 281
 QY 571 TGTCCAGAAAGAGCGCGAGGTGAGACATGTCCCGATCTCTGCTGTAGAACACAGAA 512
 Db 282 ValProProProProProGlnAsnAsnIlyProProValProSer-----ThrProArg 299
 QY 511 TGCCCTCAACAGAACACCGTCCGACAGGTAGACACCGTGTCTGTGTAACCTTGAGT 452
 Db 300 ---ProSerAlaProHisArgProHisLeuArgProPro----- 311
 QY 451 GTGTCACAAAATATATAGTGTGGGACCTCCACGCTTTTCCCTGCAGATGTTCTTCAGAG 392
 Db 312 -----ProProSerArgProGlyProProLeu---ProProSerSerSerGlyAsn 328
 QY 329 AspGluThrProArgLeuProGlnArgAsnLeuSerLeuSerSerThr----- 345
 Db 331 AGGCTTGGGCTTCTGTCTGCGGTACAGACCTTGTAACCTGTCTGCTGACAGATGA 272
 QY 346 ProProLeuProSerProGlyArgSerGlyProLeuProProProProSerGluArgPro 365
 Db 271 CGACCTTCGCGTCCGCGTGTTCACACCTGTTGTGCCAGAACCTCAGATCTTCCAC 212
 QY 366 ProProProValArgAspProProGlyArgSerGlyProLeuProProProProVal 385
 Db 211 ACAAGTGCATCTCCGCTGCGAGTCCGCGCTCACCCCTATCAGAAAGGCGGTGTG 152
 QY 386 SerArgAsnGlySer-----ThrSerArgAlaLeuPro 396
 Db 151 GCGGACGTCGCGCTCCGCGCGG---GGCTTCGCAATCTTGCCTCCCGCGGAAGCCA 95
 QY 397 AlaThrProGlnLeuProSerArgSerGlyValAspSerProArgSerGlyProArgPro 416
 Db 94 TCTCGGCTCGG---CTCCGCGGCATCGGATCCCGCGCGCGGCTTCC 47
 QY 417 ProleuProProAspArgProSerAlaGlyAlaProProProProPro 433

RESULT 12

US-09-975-719-369
 ; Sequence 369, Application US/09975719
 ; Publication No. US20030022349A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ausubel, Frederick M.
 ; APPLICANT: Rahme, Laurence G.
 ; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
 ; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
 ; FILE REFERENCE: 00786/361003
 ; CURRENT APPLICATION NUMBER: US/09/975.719
 ; PRIOR FILING DATE: 2001-10-10
 ; PRIOR APPLICATION NUMBER: US 09/199.637
 ; PRIOR FILING DATE: 1998-11-25
 ; PRIOR APPLICATION NUMBER: US 60/066,517
 ; PRIOR FILING DATE: 1997-11-25
 ; NUMBER OF SEQ ID NOS: 437
 ; SOFTWARE: FastSeq for Windows Version 4.0

[illegible]

```

Db      215 ProGlytArgSerProProGlylCySeSerAlaThrAspArgAlaIalaIaGArgSerArgLeu 234
          |||||  ||  |||||  |||  ::::|||  |||||
Qy      106 CGCGGAAGCCATCTCGGCCTCCGCCTCCGGCAGCATCGGCTC---CGCGGCCGCCCTT 50
          |||||:::  |||||  |||  |||  |||||
Db      235 ProProAArg-----MetProHisAlaHisSerArgProAlaProProGly 249
          |||||  |||||  |||||  |||||
Qy      49 CCCCCGGGCCCGCGCCGCCCGCCACGCCCG 20
          |||||  |||||  |||||  |||||
Db      250 SerProAlaValaArgAlaProAlaPro 259

RESULT 13
US-10-024-368-5
; Sequence 5, Application US/10024368
; Publication No. US20030027300A1
; GENERAL INFORMATION:
; APPLICANT: THOMPSON, Catherine C.
; TITLE OF INVENTION: HUMAN HAIRLESS GENE AND PROTEIN
; FILE REFERENCE: Thompson-20263/024435
; CURRENT FILING DATE: US/10/024,368
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/287,354
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-04-07
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/080,888
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-04-07
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1207
; TYPE: prt
; ORGANISM: Rat
US-10-024-368-5

Alignment Scores:
Pred. No.:           0.000354           Length:           1207
Score:             169.00                Matches:            135
Percent Similarity: 33.89%               Conservative:       47
Best Local Similarity: 25.14%              Mismatches:        159
Query Match:         5.74%                 Indels:            197
DB:                  15                     Gaps:              32

US-09-896-522-1 (1-1624) x US-10-024-368-5 (1-1207)
Qy      1327 TGGAAAAACTCTCTCCCACTGTGGGTTCATCTCAACAATA-----1286
          |||||:::  ::|||  |||  :::  |||  |||
Db      40 TrglnuysThrlapProvalAnsngllVlevaldylngluinProGlylThSerProGln 59
          |||||:::  ::|||  |||  :::  |||  |||
Qy      1285 -----CATKAGGCCAGCCAGTGCTTAGGGCTGTCTCTCAATTCCCACAAATAGTG 1235
          |||||:::  ::|||  |||  :::  |||  |||
Db      60 AspglYleuHIsHsglYalAlaLeuCysleudlYluProValProPheTrpArglYVal 79
          |||||:::  ::|||  |||  :::  |||  |||
Qy      1234 CCTCACATTCTCTCAAGAAGCCTC---CGAGGCTCTCGGACAT-TCTGTGGCATTTCTC 1179
          |||||:::  ::|||  |||  :::  |||  |||
Db      80 LeuSerAlaProApsSerTrpleuPProGlyPheLenuGlnglyProlysAspThrlen 99
          |||||:::  ::|||  |||  :::  |||  |||
Qy      1178 AGTGACCTAGAGGAGATCTTTAAACCGCAAC-----GAGCCTTAAGTGCGTGAANAAC 1128
          |||||:::  ::|||  |||  :::  |||  |||
Db      100 SerValValglunglYglunglYserArgnsnldYglunrglYslalaentTrpleu----- 117
          |||||:::  ::|||  |||  :::  |||  |||
Qy      1127 TCAGGAACGCTGTCAGTGCCAGCAAGTTGAAGTGAAGTACACATCTGAAGTTTCCAC 1068
          |||||:::  ::|||  |||  :::  |||  |||
Db      118 ---glySerlYsglunlYleuArg-----124
          |||||:::  ::|||  |||  :::  |||  |||
Qy      1067 TCTCGAGTGAAGAGGCGCTCGCTGTCAACTCCCTCGGGGTGGCCGAGAGAGACAGT 1008
          |||||:::  ::|||  |||  :::  |||  |||
Db      125 -----TrpLySgluaIameleuAlaHisPro-----LeuAla 135
          |||||:::  ::|||  |||  :::  |||  |||
Qy      1007 GGGTGTGGGTTGGGCGTCCCGAGGCTCACTCCCTGAACACACATGCGGCGGAGACCTG 948
          |||||:::  ::|||  |||  :::  |||  |||
Db      136 PheCySglYProAlaCySPro-----142
          |||||:::  ::|||  |||  :::  |||  |||
Qy      947 CCTTGAGGCTCGGAGCCCTCAAGTGGAGGCTGTGCTGTGAATTCCAAATGATGACCGTTTG 888
          |||||:::  ::|||  |||  :::  |||  |||

```

```

Db 143 -----ProArgTyrGlyProLeu-----11e 149
QY 887 CCAGAGTCAGCATCCAGGGTGGTCCCTGCTCAGAAAAGTCCGCTTAGCTCCGC 828
Db 150 ProGluHisSer-----SerGlyHisProLysSerProValAla-----PheArg 165
QY 827 CCATTGACCTCTCT-----CGTGCATTTTGAGATGTCACCATTCAGATG 780
Db 166 ProLeuHisCysProPheLeuGluThrLysIleLeuGluArgAlaProPhe-----183
QY 779 TCCCGATG-----TGCTGACATCAGGTGATGCGAAC-----744
Db 184 ---ThrValProThrCysLeuProProTyrIleuMetSerSerLeuProProGluArgSer 202
QY 743 -----ATATTGTCCACTCTCCCTGGGAGTATGATACA 714
Db 203 TyrAspTyrProLeuAlaProSerProTyrValTyrSerGlySerGlnProLysValPro 222
QY 713 TCGGCATACCTTTTGTGCGCA-----GGCAGAACTCTCGA-----AGGCC 672
Db 223 SerAlaPhe-SerLeuGlySerLysGlyPheTyrHisLysAspProAsnIleLeuArgPr 242
QY 671 GGCCTTCAGAAAGGTGCTACTGGCTCAGAAATCT-----638
Db 242 oAla---LysGluProLeuAlaAlaSerGlySerGlyMetLeuGlyLeuAlaProGlyG1 261
QY 637 -----GCTCCAGATCCCTCCCTCGGCGACAGTCCCGAGACT 600
Db 261 yHisLeuGlnGlnAlaCysAspAlaGluGlyProSerIleuHisGlnArgAspGlyGlu-- 280
QY 599 CTTTGAGACAGCTGACGTGCGAGTCGGTGTCCAGAAAGCGCAGCT-----GG 549
Db 281 -----ThrGlyAlaGlyArgGlnAs 288
QY 548 AACATGCCCGGATCTCTGCTGTAGAACACCAAGATCCCTCAACAGAACACAGCTCC 489
Db 288 nLeuCysProValAlaPheLeuGly-----TyrProAspThrValProArgTh 303
QY 488 G-----CAGGGTAGACACCGGTGCTGTGTAACTTG-----455
Db 303 rProTyrProSerLysProProGlyGlyCysLysSerSerHis-----410
QY 454 ---AGTGTGCACAAATCATAGATCGGACCTCA-----CCGTTTGGCCT-- 410
Db 323 oGlySerAsnSerPheGlyTyrGlnLeuGlyProProValThrProArgCysProSerPr 343
QY 409 -----CCAGATGTTCTTACAGAGTCTCTGATCAAAATTAATTAAGGATCTGATGG 354
Db 343 oGlyProProThrProProGlyGlyCysLysSerSerHis-----356
QY 353 TCAAATTTAGTCTCTTCAAGGCTTGACCTTGTGCTGCGCTCA-----305
Db 357 -----LeuProAlaArgGluGlyAs 363
QY 304 ---GACCTTTTGAACTGTCTGCTGCTCAGATGACACCTTCCGCTCCGCTGTTC 249
Db 363 pProGlyProCysArgLysCysGlnAspSer-----ProGluGlySerSerSerGlyPr 381
QY 248 ACCTCGTTCTGTCCA-----GCAACTCGATGATTTCTCA 213
Db 381 oGlyGluSerSerGluGluArgAsnLysAlaGlySerArgAlaSerProProSerHisH1 401
QY 212 CACACGGTGCATTCCTCCGCTGACAGTCCGCGCTCAACCTTATCAGAGAGGCGCTGG 153
Db 401 sThrLysLeuLysLysThrLysThrLysLysSerGlnGlnPheGluCysProGlyG1 421
QY 152 TCGGACGCTGCG-----CCTCGGCGCGGGGCTTCGAGTCTTGCGCTCCGCGGAA 99
Db 421 yCysProGlyLysGlyLysSerProAlaThrGlyLeuArgAlaLeuLysArgAlaGlySe 441
QY 98 GCGA-----TCTGGCCTCCGCTCCGCGGATCGGGTCCCGCGCGCGCGCT 51
Db 441 rProGluValGlnGlyAlaArgGlyProAlaProLys-----ArgPro-SerHisThrP 459

```

```

QY 50 TCCCGGAGCCGCGCGCCCGCCAGCGGAGTGGAGGCGACCC 4
Db 459 heProGlyThrGlyArgGlnGlyAlaArgAlaTrpGlnGluThrPro 474

RESULT 14
US-10-078-547-24
; Sequence 24, Application US/10078547
; Publication No. US20020199211A1
; GENERAL INFORMATION:
; APPLICANT: Narayanaswamy Ramesh
; APPLICANT: Miguel A. de la Fuente
; APPLICANT: Ines M. Anton
; APPLICANT: Ralf S. Geha
; TITLE OF INVENTION: WIP, A WASP-Associated Protein
; FILE REFERENCE: 1242.1022-005
; CURRENT APPLICATION NUMBER: US/10/078,547
; CURRENT FILING DATE: 2002-02-19
; PRIOR APPLICATION NUMBER: 09/599,287
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: PCT/US98/27501
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/101,457
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/068,533
; PRIOR FILING DATE: 1997-12-23
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Translated WIP ORF No. US20020199211A1 3
US-10-078-547-24

Alignment Scores:
Pred. No.: 0.00676 Length: 507
Score: 164.00 Matches: 99
Percent Similarity: 34.41% Conservative: 18
Best Local Similarity: 29.12% Mismatches: 156
Query Match: 5.57% Indels: 67
DB: 14 Gaps: 15

US-09-896-522-1 (1-1624) x US-10-078-547-24 (1-507)
QY 1036 TCCCTGGGGTGGGCGGAGAGACAGATGGGTGGGGCTCCCAAGCTCACTCC 977
Db 157 SerProGlyHisArg-----SerGlyProProGluProGlnArg 169
QY 976 CTGACACATCCCGGCGGAGACCTGCTGAGAGCTCGGACGCCCTCAGTGGGTC 917
Db 170 ---Asn-ArgMetProProProArgProAspValGlySerLysProAspSerLysProPr 188
QY 916 TGCTGTGATCTCAATATGTACCGTTTGCCAGAGTCAAGATCCCAAGGTGTCCCTG 857
Db 188 oProValProSerThrProArgProIleGlnSerSerLysHisAsnArgGlySerProPr 208
QY 856 GCTCAGAAAAGTCCGCTGTAGCTCCGCCATTTGACCTCTCTGCTGCTTGGCAGA 797
Db 208 oValProGlyGlyProArgGlnProSerPro---GlyProThrProProPhe----- 225
QY 796 TGTACCATTCAGAAATGTCTGGATGTGTGACATCAGTTGATGGCAACATATTGT 737
Db 226 -----ProGlyAsnArg--Gly 230
QY 736 CCACTCTCTGTGAGATGATCAGATCGCATCTTTGTG-----GCA 692
Db 231 ThrAlaLeuGlyGlyGlySerLysLysArgLysSerProLeuSerSerSerSerProPheSer 250
QY 691 GCGACAACTCTCGAAGCGCGGCTTCACGAAGGTGTACTGCGTCAAGAACTGCTCA 632
Db 251 AsnArgProProLeuProProThrProSerArgAlaLeuAspLysProProProPro 270

```

```

QY      631 GGNCCCTCCCTGGCGCAGSTCCCGAGAACTTTCAGAGACAGCTGAGCTCGGAGTCCG 572
Db      271 ProProProValGlyAsnHrPro-----SerLeHisArgGlyAla 284
QY      571 TGTCCAGAGAGAGCCAGGTGGACATGTCGCCGATCTCTGGCTGTAGAACACCAAGA 512
Db      285 ValProProProProProGlnManAsnLysProProValProSer-----ThrProArg 302
QY      511 TGCCTTCAAAAGAACCAAGTCCCGCAGGGTAGAACACACCGCTGTCTGTGAACCTTGAGT 452
Db      303 ---ProSerAlaProHisArgProHisLeuArgProPro----- 314
QY      451 GTGTCAAAAATATATATGTGGCAGCATCTCCACGTTTTCGCCACAGATCTTTCAGAG 392
Db      315 -----ProProSerArgProGlyProProProLeu---ProProSerSerArgGlyAsn 331
QY      391 TCCTGTGCATCAATATATTATCAAGGACATGTGATGTCAAAATTGTACTGTCTTTC 332
Db      332 AspGluThrProArgLeuProGlnArgAsnLeuSerLeuSerSerThr----- 348
QY      331 AGGCTTGGCTTTCGTCTGCGTCAAGACCTTGT---AGAACTGTCTGTGCTCAGGA 275
Db      349 ProProLeuProSerProGlyArgSerGlyProLeuProProProValProSerGluArg 368
QY      274 TGACCACTTCGCGCTCGCTGTTCACCTCTGTCTGTCCAGCAACTGCATATCTTCT 215
Db      369 ProProProProValArgAsnProProGlyArgSerGlyProLeuProProProPro 388
QY      214 CACACAGGTGCAGTCTCCGCTGGCAGTGCAGCGCGCTCACCCCTATCAGGAAGGCGGCT 155
Db      389 ValSerArgAsnLysLeu-----ThrSerAlaLeu 399
QY      154 GGTGGAGAGCGTGGCGCTCGGCGCGG---GGCTCTTCGAGTTCGCTCCGCGCGAAG 98
Db      400 ProAlaThrProGlnLeuProSerArgSerGlyValAsnSerProArgSerGlyProArg 419
QY      97 CCATCTCCGCGCTCG---CTCCCGCGCATGGGTCCCGCGCGCGCGCCCTTCCC 47
Db      420 ProProLeuProProAsnArgProSerAlaGlyAlaProProProProPro 437

RESULT 15
US-10-017-161-1350
; Sequence 1350, Application US/10017161
; Publication No. US20030143668A1
; GENERAL INFORMATION:
; APPLICANT: SUMA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: NOVEL G PROTEIN-COUPLED RECEPTORS
; FILE REFERENCE: 084335/0152
; CURRENT APPLICATION NUMBER: US/10/017,161
; PRIOR FILING DATE: 2002-12-18
; PRIOR APPLICATION NUMBER: JP 2001/246789
; NUMBER OF SEQ ID NOS: 2430
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1350
; LENGTH: 737
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (2)..(35)
; OTHER INFORMATION: Variable amino acid
US-10-017-161-1350

Alignment Scores:
Pred. No.: 0.000744 Length: 737
Score: 164.00 Matches: 148
Percent Similarity: 33.16% Conservative: 44
Best Local Similarity: 25.56% Mismatches: 209

```

```

Query Match: 5.57% Indels: 178
DB: 12 Gaps: 29
US-09-896-522-1 (1-1624) x US-10-017-161-1350 (1-737)

QY      1339 ACCAGATCAGATGTGAAAAAATCTCTCCACTGTGGTTCATCTGCAACAAA-----A 1286
Db      172 ThrProGlnSerThrLysProLeuSerProGlyAsnProSer--ValArgLysProLeuS 191
QY      1285 CATCAGGCGCAGCGAGTGTCT-----AGCTGTCTCTCAATTCCCAATAATGTG 1235
Db      191 exProGlySerProLeuSerProArgLysProLeuSerProGlyAsnProSerValProG 211
QY      1234 CCTCAGATCTCCACAGAACGCT-----CCGAGGCTCTCGACATTCG 1190
Db      211 LThrProGlnSerGlyAsnProSerValProGlyAsnProSerValProGlyAsn---- 229
QY      1189 TGGCATTTTCTCAGTGCATGACGAGATCTTTAAACGCAACGACGCTAAGTGGCTGA 1130
Db      230 -----ProSerValProGlyThrProGlnSerArgLysPro-----ProG 243
QY      1129 CCTCAGGAAGCGCTGTCACTGTTCACGAGTTGAGTCTGAGTACACA-----TCTGA 1076
Db      243 InsSerGlyAsnPro--SerValProGlyThrProGlnSerGln-GluThrProGlnSerAr 262
QY      1075 GTTTCACCTCTCAGTGAAGGAGGAGGCGCTGCTGTAACAATCCCTCGGG----- 1027
Db      262 GlySProLeuSerProArgAsnProPheSerProArgAsnProLeuGlyProGlyAsnBr 282
QY      1026 -----TGCGCCGAGAGAAAGCATGTGGGTGTGGTGGCGTCC 989
Db      282 LeuSerGlnGluThrProGlnSerArgGluProGlnInsSerGlyAsnProSerValPr 302
QY      988 CAGG----- 985
Db      302 oArgAsnProLeuSerProGlyAsnProSerAlaProGlyThrProGlnSerLeuGluPr 322
QY      984 -----CTCAGTCCCTGAACACACATCCGCGGCGGAGACCGGCCCT 944
Db      322 LeuSerProGlyLysProLeuSerProGlyAsnProPheAsnProGlyAsnProSerVa 342
QY      943 GAGGCTC----- 937
Db      342 LArgLysProLeuSerProGlyAsnProLeuSerProGlyLysProLeuSerProArgAs 362
QY      936 -----GGCAGCCCTCAGTGGGCTGCTGCTGACTTCCAAA-----T 899
Db      362 nProSerValProGlyAsnProSerValProGlyThrPro-GlnSerArgLysProPheS 382
QY      898 GTGACCGTTTGCAGAGGTACAGATCCACGGGTGTCCTCGCTCAGAAAAGTCCGCT 839
Db      382 exProArgAsnProSerValSerGlyThrProSerValProGlyThrProGlnSerArgL 402
QY      838 TGTAGCTCCGCCAATTGACCGCT-----CTTC 812
Db      402 yeProLeuSerProGlyAsnProPheGlyProGlyAsnProLeuValProGlyAsnProS 422
QY      811 GGTGCATTTCAGATGTACCAATTCAGAAATGCTCTGATGTGTGCAGATCAGGTGA 752
Db      422 exValArgLysProLeuSerProGlyAsnPro-----L 433
QY      751 TGGCAACCATATTGTCCACTCTCGTGGATATATACATCGGATATCTTTGTGGCA 692
Db      433 euser-ProGlyLysProLeuSerProGlyAsnProSerValArgLysProLeuSerPro 452
QY      691 GGCAGA-----ACTCTCAGAGCGCGCTTACAGAAAGGTGTGTACTCGTCAGATCT 638
Db      453 GlyLysProLeuSerProArgLysProLeuSerProGlyAsnProGlnSerArgLysPro 472
QY      637 ---GCTCAGGT-----CCCTTCCTGGCGCAGTCCCGGAGACTTTTCAG 593
Db      473 LeuSerProGlyLysProLeuSerProGlyLysProLeuSerProGlyAsn----- 489

```

```
QY 592 ACAGCTGACGTCGAGTCGCTGCACGAAGAGCGCAGGTGAACAATGTCCTCGATCT 533
Db 490 -----ProleuserProglyLysProleuCysProglyAsn 501
QY 532 CCTGCTGTAGACACCAAGATGCCCT-----CAAACGAACCAAGTCGCG 488
Db 502 ProseValArgLysProleuserProglyLysThrProGlnSerArgGluProleuser 521
QY 487 CAGGGTAGACCA-----CCGTGCTCTGTAACTTGAGTGATGTACAA 443
Db 522 ProglyLysProleuserProglyAsnProGlnSerArgGluProleuserProglyLys 541
QY 442 -----AATCATAGCTCGGCACTTCACCGTTT 416
Db 542 ProleuserProglyLysProleuserProglyAsnProleuserProglyLysProleu 561
QY 415 TGCCCTCCAGATGTTCTTCAGAGTCCTGTCATCAATCATTCATAAGGCACTTGAT 356
Db 562 CysProglyAsnProseValArgLysProleuserPro-----GlyLysProleuser 579
QY 355 GGTCAAAATTGTAAGTCCTTCAAGGCTTGACCTTGCTGCTGCGTCAGACCTTGT 296
Db 580 ProglyAsnProseValProglyAsnProseValProglyThrProseValProgly 599
QY 295 AGA---ACCTGTCCTGGCTCAGATGACCA-----CCTTCGCT 260
Db 600 AsnProseValProglyThr-----ProGlnSerArgAsnProseValProglyAsn 617
QY 259 GCCGCTGTTCCACTCGT---TCTGTCCAGCACTCCATGATCTTCTCACACAGGTCG 203
Db 618 ProseValProglyAsnProseValProglyThrPro-----GlnSer 632
QY 202 ACTTCCCGCTGGCAGTGCAGCGCGCTCAACCCTATCAGGAAGCGCGCTGTGCGACGCT 143
Db 633 GlnGluProProGlnSerArgGluProProGlnSerArgGluThrProAlaValProgly 652
QY 142 CGGCTCCGGGCGGGGCTCTCGCAGTCTTGCTCCGCGGAGCAATCTCGGCTCCG 83
Db 653 AsnPro-----SerValArgLysProleuserProglyLysProleu 666
QY 82 CTCCCGCGC-----ATCGGGTCCCGCGCGCCGCT---CCCGGGCGCGG 38
Db 667 SerProArgLysProleuserProglyThrProseValProValThrProGlnSerArg 686
QY 37 -----CGCGCCCGCCAGCGCGAGGTGCGAGGCGACCCCA 2
Db 687 GluThrProGlnSerArgGluProProGlnSerArgGluThrPro 701
```

Search completed: November 25, 2003, 07:50:17
Job time : 187.943 secs


```
QY 214 GAGAGATCATGAGTGTCTGGGACAGAACGAGGTGGAACAAGCGGACCGGAAAGGTGTC 273
Db 41 Glnuyllelmetcgluleuenglglinsnglvalgluglnarglnarglinsvalval 60
QY 274 ATCTGAGCCAGGACAGGTCTCAAGAGTCTTGACGGCAGAGCAAGAGGCAAGGCGCTTG 333
Db 61 lIleuSerGlnsParghneTyrLysValleuThrAgluGlnLysAlaLysAlaLeu 80
QY 334 AAAGACATGACATTTTGAACCATCCAGATGCTTTGATATGATATTTGATGACAGACT 393
Db 81 LyselYlInTyrAsnPhenPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
QY 394 CTGAAGACACTCGTGGAGGGCAAAACGGTGGAGGTGCCACTATGATTTTGTGACACAC 453
Db 101 LeuLysAsnIleValGlnGlyLysThrValGlnValProThrTyrAspPheValThrHis 120
QY 454 TCAAGGTACAGAGACAGGAGTGTCTACCGCTGCGAGCGTGTCTGTTTGAAGGCATC 513
Db 121 SerArgLeuProGlnThrThrValValTyrProAlaAspValValLeuPheGlnGlyIle 140
QY 514 TTGGTGTCTTACAGCCAGAGATCCGGACATGTTCCACCTGCGCTCTTGTGACACC 573
Db 141 LeuValPheTyrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
QY 574 GACTCCGACGTCCAGGCTGTCTCGAAGAGTTCTCCGGACGTGCCCGAGGAGGACCTG 633
Db 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
QY 634 GAGAGATTTGACGACGATACACACCTTCGTGAAGCGGCGCTTCAGAGAGTTGTGCTG 693
Db 181 GlnGlnIleLeuThrGlnTyrThrThrPheValLysProAlaPheGlnGlnLysPheCysLeu 200
QY 694 CCGACAAAGAATATGCCATGTGATCATCCACAGAGAGTGACAAATATGTTGCCATC 753
Db 201 ProThrLysTyrLysThrAlaAspValIleIleProArgGlyValaAspMetValAlaIle 220
QY 754 AACCTGATCTGTCGACGACATCCAGACATTTGAAATGTGACATCTCCAAATGGCACCGA 813
Db 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlnLysPheCysLysThrPheArg 240
QY 814 GAGAGGTCCAAATGGGCGAGCTACAGCGGACCTTTCTGAGCAGGAGGACCACTGGG 873
Db 241 GlnGlySerAsnGlnLysArgSerTyrLysArgThrPheSerGlnProGlyAspHisProGly 260

RESULT 2
US-09-536-647-3
; Sequence 3, Application US/09536647
; Patent No. 6579708
; GENERAL INFORMATION:
; APPLICANT: Ho, Yen Sen
; APPLICANT: Johnson, Randall
; TITLE OF INVENTION: No. 6579708el Human Uridine Kinase
; FILE REFERENCE: GP50020
; CURRENT APPLICATION NUMBER: US/09/536,647
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 3
; LENGTH: 277
; TYPE: PRT
; ORGANISM: Human
US-09-536-647-3

Alignment Scores:
Pred. No.: 4,37e-109 Length: 277
Score: 1325.00 Matches: 254
Percent Similarity: 95.67% Conservative: 11
Best Local Similarity: 91.70% Mismatches: 12
Query Match: 44.66% Indels: 0
Gaps: 0
US-09-896-522-1 (1-1624) X US-09-536-647-3 (1-277)
```

```
QY 94 ATGCTTCGGCCGGAGGCGCAAGACTCCGAGACCCCGCCCGGAGGCCGACCGCTCCGAC 153
Db 1 MetAlaSerAlaGlyGlyLysSerGlnSerAlaAlaProGlnLysAlaAspArgProGln 20
QY 154 CAGCGGCGCTCTCTGTAAGGGGTGAGCGGCGGACCTCCAGCGGGAAGTGCACCGTGTGT 213
Db 21 ProArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrAlaCys 40
QY 214 GAGAGATCATGAGTGTCTGGGACAGAACGAGGTGAAACAGCGGACCGGAAAGTGTGTC 273
Db 41 GlnuyllelmetcgluleuenglglinsnglvalasparGArgGlnArgLysLeuVal 60
QY 274 ATCTGAGCCAGGACAGGTCTCAAGAGTCTTGACGGCAGAGCAAGAGGCAAGGCGCTTG 333
Db 61 lIleuSerGlnAspCysPheTyrLysValLeuThrAlaGlnGlnLysAlaLysAlaLeu 80
QY 334 AAAGACATGACATTTTGAACCATCCAGATGCTTTGATATGATATTTGATGACAGACT 393
Db 81 LysGlyGlnTyrAsnPhenPheAspHisProAspAlaPheAspAsnAspLeuMetHisArgThr 100
QY 394 CTGAAGACACTCGTGGAGGGCAAAACGGTGGAGGTGCCACTATGATTTTGTGACACAC 453
Db 101 LeuLysAsnIleValGlnGlyLysThrValGlnValProThrTyrAspPheValThrHis 120
QY 454 TCAAGGTACAGAGACAGGAGTGTCTACCGCTGCGAGCGTGTCTGTTTGAAGGCATC 513
Db 121 SerArgLeuProGlnThrThrValValTyrProAlaAspValValLeuPheGlnGlyIle 140
QY 514 TTGGTGTCTTACAGCCAGAGATCCGGACATGTTCCACCTGCGCTCTTGTGACACC 573
Db 141 LeuValPheTyrThrGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
QY 574 GACTCCGACGTCCAGGCTGTCTCGAAGAGTTCTCCGGACGTGCCCGAGGAGGACCTG 633
Db 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValGlnArgGlyArgAspLeu 180
QY 634 GAGCAGATTTGACGACGATACACACCTTCGTGAAGCGGCGCTTCAGAGAGTTGTGCTG 693
Db 181 GlnGlnIleLeuThrGlnTyrThrAlaPheValLysProAlaPheGlnGlnLysPheCysLeu 200
QY 694 CCGACAAAGAATATGCCATGTGATCATCCACAGAGAGTGACAAATATGTTGCCATC 753
Db 201 ProThrLysTyrLysThrAlaAspValIleIleProArgGlyValaAspMetValAlaIle 220
QY 754 AACCTGATCTGTCGACGACATCCAGACATTTGAAATGTGACATCTCCAAATGGCACCGA 813
Db 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlnLysPheCysLysThrPheArg 240
QY 814 GAGAGGTCCAAATGGGCGAGCTACAGCGGACCTTTCTGAGCAGGAGGACCACTGGG 873
Db 241 GlnGlyProAsnGlnLysArgSerHisLysArgThrPheProGlnProGlyAspHisProGly 260
QY 874 ATCTGACCTCTGCAACCGTCACTTTGATGATGATGATGATGATGATGATGATGATGAT 924
Db 261 ValLeuAlaThrGlyLysArgSerHisLeuGlnSerSerArgProHis 277

RESULT 3
US-09-134-001C-3618
; Sequence 3618, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
; FILE REFERENCE: GTC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR APPLICATION NUMBER: 1998-08-13
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
```

```

: SEQ ID NO 3618
: LENGTH: 216
: TYPE: PRT
: ORGANISM: Staphylococcus epidermidis
: US-09-134-001C-3618

Alignment Scores:
Pred. No.: 9.69e-32 Length: 216
Score: 451.00 Matches: 91
Percent Similarity: 63.68% Conservative: 44
Best Local Similarity: 42.92% Mismatches: 61
Query Match: 15.20% Indels: 16
Gaps: 4

US-09-896-522-1 (1-1624) x US-09-134-001C-3618 (1-216)
QY 166 CTGATAGGGGTGAGCGCGCGGCACTGCCAGCGGAAAGTGCAGCCGTGTGAGAAATCATG 225
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 15 lTlleglylTlealaglyglYserclYserclYserclYserlthrrlvalTlrAsnAlalelmet 34
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 226 GAGTTGCTGGGACAGAACGAGGTGGAACAGCGGACGCGGAGGAGGTGCATCTTGAGCAG 285
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 35 LysAsnleu-----GluglYhIsSerValAlaleuValaglIn 47
   ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 286 GACAGGTTCTTCAAG-----GTCTGACGCGACAGACAGAAAGGCCAGGCTTGAAA 336
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 48 AsprYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYr 65
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 337 GGAACAGTACATTTTGACATCCAGATGCTTTGTATATGATATGATATGATATGATATGAT 396
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 66 -----AntYrAspHisrProPheAlaPheAspHisAsnLeuLeuIleHisAsnLeu 82
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 397 AAGAATCTGTGAGGGGCAAAACGGTGGAGGTGCGGACCTATGATTTTGTGACACACTCA 456
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 83 LysAspLeuArGAsnGlyLyLeProValGlulValProThrYrArYrYrYrYrYrYrYrYrYr 102
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 457 AGGTTACAGAGACACCGGTGTCTACCTTCGCGACCGGTGCTGTCTTTGAGGCGATCTTG 516
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 103 ArGerYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYr 122
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 517 GTGTTCACAGCCAGAGATCGGGAGACATGTTCCACCTCGGCGCTTCGTGGAACCGAC 576
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 123 AlAlaLeuGlAsnAsnAntHrLeuArYrAspMetCetAspValYrIleYrValAspThrAsp 142
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 577 TCCGACGTCAGGCTGTCTCGAAGATTCTCCGGACGTCGCC---CGAGGAGGAGCACTG 633
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 143 AlAspLeuArGlIleLeuArGArYrLeuThrArGAspThrYrYrYrYrYrYrYrYrYrYrYrYr 162
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 634 GAGCAGATTCTGACGCGACGTACCACTTCGTGAAGCCGGCTTCGAGAGATTGTCCTG 693
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 163 GluSerValIleAsnGlnYrYrLeuAsnValYAlArGProMetHisGluglnPheIleGlu 182
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 694 CCGACAAAGAAAGTATGCCGATGATATATCCAGAGAGAGTGCACATATAGCTTGCCATC 753
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 183 ProThrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYr 202
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
QY 754 AACCTGATCGGACGACATCCAGACATTTGAT 789
   |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db 203 AspIleMetHrThrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYrYr 214

RESULT 4
US-09-107-532A-5024
: Sequence 5024, Application US/09107532A
: Patent No. 6583275
: GENERAL INFORMATION:
: APPLICANT: Lynn A Doucette-Stamm and David Bush
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
: ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
: NUMBER OF SEQUENCES: 7310
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: GENOME THERAPEUTICS CORPORATION
: STREET: 100 Beaver Street
: CITY: Waltham

```

```

1 STATE: Massachusetts
2 COUNTRY: USA
3 ZIP: 02354
4
5 COMPUTER READABLE FORM:
6 MEDIUM TYPE: CD/ROM ISO9660
7 COMPUTER: PC
8 OPERATING SYSTEM: <Unknown>
9 SOFTWARE: ASCII
10
11 CURRENT APPLICATION DATA:
12 APPLICATION NUMBER: US/09/107,532A
13 FILING DATE: 30-Jun-1998
14
15 PRIOR APPLICATION DATA:
16 APPLICATION NUMBER: 60/085,598
17 FILING DATE: 14 May 1998
18 APPLICATION NUMBER: 60/051571
19 FILING DATE: July 2, 1997
20
21 ATTORNEY/AGENT INFORMATION:
22 NAME: Ariniello, Pamela Deneke
23 REGISTRATION NUMBER: 40,489
24 REFERENCE/DOCKET NUMBER: GTC-012
25
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE: (781)893-5007
28 TELEFAX: (781)893-8277
29
30 INFORMATION FOR SEQ ID NO: 5024:
31
32 SEQUENCE CHARACTERISTICS:
33 LENGTH: 210 amino acids
34 TYPE: amino acid
35 TOPOLOGY: linear
36 MOLECULE TYPE: protein
37 HYPOTHEICAL: YES
38 ORIGINAL SOURCE:
39 ORGANISM: Enterococcus faecium
40
41 FEATURE:
42 NAME/KEY: misc feature
43 LOCATION: (B) LOCATION 1..210
44
45 SEQUENCE DESCRIPTION: SEQ ID NO: 5024:
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
```

```
Db 114 GlyIleuLeuIleuGluAspGluArgLeuArgSerLeuMetAspIleValTyrVal 133
|||
QY 568 GACACCGCATCCGACGTGAGCTGTCTGAAAGCTTCTCGGGACGTG---CGCCGAGG 624
|||
Db 134 AspThrAspAspAspIleArgIleIleArgArgIleuArgSerMetGluGluArgGly 153
|||
QY 625 AGGACCTGGAGCAGATTCTGACCGACGATACCACTTCCTGGAAGCCGGCTTCGAGAG 684
|||
Db 154 ArgThrLeuAspSerValIleGluGlnTyrLeuThrValValValValProMetTyrHisGln 173
|||
QY 685 TTTCGCTGCGCCGACAAAGATATGCGATGTGATCATCCACAGAGAGTGACATATG 744
|||
Db 174 PheIleGluProThrLysArgTyrHisAspIleIleValProGluGluArgLysAsnHis 193
|||
QY 745 GTTGCCATCAACGTGATGTGACACACATCCAGACATTTGAT 789
|||
Db 194 ValAlaIleAspLeuIleThrThrValAlaSerPheLeuAsn 208
|||

RESULT 5
US-09-198-452A-783
; Sequence 783, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Griffois, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, preve
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198-452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 783
; LENGTH: 222
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-783

Alignment Scores:
Pred. No.: 2,79e-28 Length: 222
Score: 412.00 Matches: 84
Percent Similarity: 63.01% Conservative: 54
Best Local Similarity: 38.36% Mismatches: 59
Query Match: 13.89% Indels: 22
DB: Gaps: 6

US-09-896-522-1 (1-1624) x US-09-198-452A-783 (1-222)
QY 166 CTGATAGGGGTGACGGCGGACACTGCACGCGGAAGTGCACCGTGTGAGAAAGATCATG 225
|||
Db 8 IleIleGlyIleThrGlySerGlyAlaGlyLysThrThrLeuThrGlnAsnIleLys 27
|||
QY 226 GAGTGTGCGGACGAACGAGAGTGAAACGCGGACGGAGAGTGTGATCTGAGGAC 285
|||
Db 28 GluIlePheGlyGluAsp-----ValSerValIleCysGln 39
|||
QY 286 GACAGGTTCTCAAG-----GTCTGACGGCAGACGAGAGGCCAAGGCTTGA 336
|||
Db 40 AspAsnTyrLysAspArgSerHisTyrThrProGluGluArgAlaAsnLeuIle--- 58
|||
QY 337 GGAAGTACAAATTTGACCATCGATGCTTGTATATGATTGTGATGACAGGACTCTG 396
|||
Db 59 -----TTPAspHisPheAspAlaPheAspAsnAspLeuIleSerAspIle 74
|||
QY 397 AAGAATCTGTGAGGGCAAAACGTGAGAGGTGCCGACTATGATTTTGTGACACTCA 456
|||
Db 75 LysArgLeuLysAsnAsnGluIleValGlnAlaProValPheAspPheValLeuGlyAsn 94
|||
QY 457 AGG---TTACAGAGACACAGCGTGTCTACCTTCGAGACGTGTCTCTTTGAGGAGCATC 513
|||
Db 95 ArgSerLysThrGluIleGluThrIleTyrProSerLysValIleLeuValGluGlyIle 114
|||
QY 514 TTGGTGTCTTACAGCCAGAGATCCGGGACATGTTTCCACCTGCGCCTTCTGTGACACC 573
|||
```

```
Db 115 LeuValPheGluAsnGlnGlnIleuAspLeuMetAspIleArgIlePheValAspThr 134
|||
QY 574 GACTCCGACGTACAGGCTGTCTCGAAGATTCTCCGGACAGTGGC---CGAGGAGGAC 630
|||
Db 135 AspAlaAspGluArgIleLeuArgMetValArgAspValGlnGlnIleGlnLysAspSer 154
|||
QY 631 CTGGACAGATTCTGACCGAGATACACACCTTCGTGAACCGGGCTTCGAGAGTCTCG 690
|||
Db 155 ValAspCysIleMetSerArgTyrLeuSerMetValLysProMetHisGlnLysPheIle 174
|||
QY 691 CTGCCCAAAAGATATGCCGATGTGATCATCCACAGAGAGTGACATATGTTGCC 750
|||
Db 175 GluProThrArgLysTyrHisAspIleIleValHisGlyAsnTyrArgGlnMetVal 194
|||
QY 751 ATGACCTGATCGTCGAG-----CACATCCAGACATTTGATGATGAC 795
|||
Db 195 ThrAsnIleuSerGlnLysIleLysAsnHisIleGlnAsnAlaLeuGluSerAsp 213
|||

RESULT 6
US-09-252-991A-30843
; Sequence 30843, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfeld et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196-136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 30843
; LENGTH: 663
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-30843

Alignment Scores:
Pred. No.: 1,07e-09 Length: 663
Score: 204.50 Matches: 147
Percent Similarity: 29.20% Conservative: 46
Best Local Similarity: 22.24% Mismatches: 183
Query Match: 6.94% Indels: 285
DB: Gaps: 37

US-09-896-522-1 (1-1624) x US-09-252-991A-30843 (1-663)
QY 1619 CCGGCAAAAGCTCG-----GTGTGATGCTGGGTGCTTCCAGGGCAT 1578
|||
Db 132 ProAlaSerAlaTyrPrtPrtLeuAlaProAlaArgCysCysTyr-----ProAlaPro 149
|||
QY 1577 CCGCTCGCTTCCTGCGCTTTAAGGACCAAGAGGGGGAAC----- 1533
|||
Db 150 ProAlaProAlaSerAlaGlyArgLacCysCysAlaSerProAsnArgArgArgGlu 169
|||
QY 1532 -----ATCCCT 1527
|||
Db 170 ProTyrProProSerProTyrAlaSerArgAlaGlyProAlaSerCysGlyArgProPro 189
|||
QY 1526 CAGTGGCTCCCATATCCGTGAGGACAGCTGGAATCGAAMACCACTGGGTATGAAATT 1467
|||
Db 190 AlaCysSerProValAlaThrAlaProThrAlaThrCysSerProProSerAlaArgSer 209
|||
QY 1466 AGTTTGTAGGCCAACACAGCCTGTGTATGAACACAAAAACAAATGAGG--- 1410
|||
Db 210 AlaTyr-----LysProCysAlaCysAlaLysAlaValGlySerAlaArgSer 225
|||
QY 1409 -----ACATGTAACATGTAAAAAAG 1389
|||
```

```

Db      226  Prohlaaenla1atProhlaa1aSerAla1atPrcyProthrCysaenSerAlaArgPro 245
Qy      1388  AAGACTTGTTCACAAAACCTTTGAGTATGTGTGAGTGTGTAGAACACAGATCAGA 1329
Db      246  SerAla1a1a1aSer-----250
Qy      1328  CTGGAAAAAACTCTCCCACTGGGTTCACTGTCAACAA-----ACATCAGG 1278
Db      251  -----ThrlaArgaSerProserThrlaThly-----534
Qy      1277  CAGCGACTGTCTAGCTGTCTCTCAATTCCCAATATATGTGCTGCATCTCTCAAG 1218
Db      261  ArgPro-----AlaAlaCysProthrThrlaThr 269
Qy      1217  AAGCTCCCAAGGCTTCTGCACATCTGTGGCATTTCTCAGTCAGCTAAGGAGATCTT 1158
Db      270  Prohlaaenla1aProhlaa1a1a1a1a-----280
Qy      1157  AACCGCAACGAGCTAGTGGCTGAAAACCTCA---GGAACGCTGTCAAGTCCAGCA 1101
Db      281  -----AsnTrpProtpTrpLysAlaProserProhlaeProProserThrla 298
Qy      1100  AGTTGACTGTAGTACACATCTGAGTTTC---CACTCTGAGTGAAGAGGCTCTG 1047
Db      299  -----ThrlaGleuAla1aCysGlyH1aAsp---GlyAspAlaGlyArg 313
Qy      1046  CTGCTAACACTCCCTCG-----GGTGGCCGAG---AGGA 1014
Db      313  laSerGlnTrpTrpArgAspThrlaSerProCysValProAsnCysValAsnTrpArg 333
Qy      1013  AGCAGTGGTGTGGGTGGGCGTCCCAAGGCTCAGTCCCTGAACACACATGCGGGGGA 954
Db      333  rgaGllleProtpSerAlaValAla1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 352
Qy      953  GACCTGCTGAGGCTGCGAGGAGCCCTCAGTGGGCTGTCTGTGATCCCAATGTGAC 894
Db      352  etProAla---GlyProArgProserCysThrlaValAla1a1a1a1a1a1a1a1a1a 368
Qy      893  CGTTTGGCAAGGAGTACATCCCAAGGCTGTCTGTGCTCAGAAAAGTCCGTTGAG 834
Db      369  -----AlaArgGlyGlyTrpAla1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 382
Qy      833  CTTCCGCCATTTGACCCCTCGGTGCGATTTGAGATGACATTCAGAAATCTCTGG 774
Db      382  rAlaGlyArg-----LeuArgArgValArgArg-----392
Qy      773  ATGTGCTGACGATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 714
Db      393  -----SerAla1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 401
Qy      713  TCGGCATACTCTTTTGTGGCAGGACGA---ACTCT---CGAAGCGCGC 669
Db      402  -----ProhlaGlyArgAlaThrlaProhlaSerAlaArgArgArg 416
Qy      668  TTGACGAGG-----TGTGTACTGTGCTC 645
Db      416  rAlaSerArgGlyArgProhlaThrlaAla1a1a1a1a1a1a1a1a1a1a1a1a1a1a 436
Qy      644  AGAA-----TCTGCTCAGGTCCTCCCTC-----620
Db      436  aArgArgThrlaSerSerAlaProhla1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 456
Qy      619  -----GGCGACGTCCTCCGGAACCTTTGACAGACGCTTACG 582
Db      456  gCysHlaProvalArgLysArgGlyAlaGlyProhlaSerThrlaArgSerArgArg 476
Qy      581  TCGGAGTCGGTGTCA-----CGAAGAGCGCGAGGTGGAAC 546
Db      476  gTrpAla1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 496
Qy      545  ATGT-----CCGAGATCTCTGCTGTAGAACACCAAGATCCCTGAACAGAACAGC 492
Db      496  aCysTrpProProhlaProhlaProhlaArgArgThrlaAla1a1a1a1a1a1a1a1a 512

```

```

Qy      491  TCCGAGGGGTAGACCAACCGTGTCTGTGTAACCTTGAGTGTGCACAAATCATAGTTC 432
Db      513  -----ArgSerAlaArgTh 517
Qy      431  GGCACCTCCACCGTTTGGCCCTCCAGATGTTCTTTCAGATCCTGTGATCAATCATTA 372
Db      517  rAlaProArgProhla1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 534
Qy      371  TCAAGGATCTGGATGTCAAAATTTACTGTCTTCAAGGCTTGGCTTGTCTCT 312
Db      535  -----TrpPro-----536
Qy      311  GCGGTACGACCTTGTAGAACCTTGCT---GGCTCAGATGACCAACTTCCGCTGCCG 255
Db      537  -----GlyProAlaLeuAlaCysProhla1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 554
Qy      254  TGTTCGA-----CCTGTCTGTCCAGCACTCATGATCTTTCACACAGGTC 204
Db      554  oAlaProhlaProhlaProhla1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 567
Qy      203  GACTCCCGCTGAGTGCAGTCCGCGCTCACCCCTATCAGAAAGGCGCGTGTGCGAGCG 144
Db      568  -----ArgArgArgPro-----G1 572
Qy      143  TCGGCTCCGCGCGCGGGGCTTCGCACTTCGCTCCG-----CGAAGCATCTG 90
Db      572  yArgSerProhlaThrlaProhla1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 592
Qy      89  GCCTCCGCTCCGCGATCGGATGCTCCG---CGCCGCGCTTCCCGGCGCGCGCGCG 33
Db      592  gProProhla1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 612
Qy      32  C-----CGCCAGCGCCGAGTGC 14
Db      612  lArgTrpProtpArgSerAlaArgCysSerProArgArgProThrlaProhlaSer 630

RESULT 7
US-09-252-991A-25394
; Sequence 25394, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252, 991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074, 788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094, 190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25394
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25394

Alignment Scores:
Pred. No.: 1,12e-09 Length: 369
Score: 203.00 Matches: 130
Percent Similarity: 35.76% Conservative: 37
Best Local Similarity: 27.84% Mismatches: 158
Query Match: 6.89% Indels: 142
DB: 4 Gaps: 27

US-09-896-522-1 (1-1624) x US-09-252-991A-25394 (1-369)
Qy      1284  ATGAGCGCAGCGCAGTCT-----TAGCTGTCTCTCAATTTCCCAATTAATG 1237
Db      4  LeuArgProhla1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a1a 23

```

```

QY 1236 TGCCCTCATCTCCACAGAAAGCCCTCCAGGCTTCCTGCACATTCCTGTGGATTTCTCAG 1177
Db      :::::
QY 24 GlyAlaArgThrMetProCysGlySerArgAlaGlyArgGlyLysArgAlaArgHis 43
Db      :::::
QY 1176 TGACCTAAGAGGATTTAAACCGCAGACGCTTAAGTGGCTGAAAACCTCAGAGAGCC 1117
Db      :::::
QY 44 CysProAlaArgProAlaArgProSerArgSerAlaAlaGlySer--ArgAlaAla 62
Db      :::::
QY 1116 TGTCAGTGTCCCGCAAGTTAGTCTGAGTGACACATCTGAGTTTCCACTCCTGAGTGAG 1057
Db      :::::
QY 63 SerArgCysProSer----- 67
Db      :::::
QY 1056 GAAAGCCTCGCTCTAACACTCCCTGGGTC-----GCCAGAGAGAGCAGTGG 1006
Db      :::::
QY 68 -----AlaAlaThrPro-----CysHisProThrLeuArgArgArgThr 81
Db      :::::
QY 1005 GTGTGGTGGGGGCTCCAGGCTCAGTCCCTGAGACACATCCGGGGGAGAGACTGCC 946
Db      :::::
QY 82 --ProGlySerArgProProArgSerSer-----ThrGlyProGlyArg-----P 97
Db      :::::
QY 945 CTGAGGCTCGGACGCGCTCAGTGGGGTCTGCTGTCGTCGACCTCAATGTGACCGTTGCC 886
Db      :::::
QY 97 roProArgArgArgProProArgArgSerAlaAlaAlaSerProArg-----H 113
Db      :::::
QY 885 AGAGGTGACATCCCGAGGTGGTCCCTGCTCAGAAAAGTCCGCTTGAGTCCGCC 826
Db      :::::
QY 113 iSAgtThrSerAlaProArgGlyPro-----GlyArgArgThrProAlaSerSer--- 129
Db      :::::
QY 825 ATTGAGCCCTCCTCGGTGCCATTGACAGATGTCACCATCTGAGATGCTCTGGATGCTG 766
Db      :::::
QY 130 -----SerGluArgProGlyThrAla- 136
Db      :::::
QY 765 CACGATCAGTTGATGCAACCATTTGTCCACTCCTCGTGGATGATCATCGGCATA 706
Db      :::::
QY 137 -----ProAlaProPro-----ArgArgT 143
Db      :::::
QY 705 CT-----TCTTGTGGCGAGGAGAACTCTCGAAGCGCG 670
Db      :::::
QY 143 hrAlaProGlyTyrProGlyProSerAlaAlaGlyAlaGlyArgProAlaArgArgProAla 163
Db      :::::
QY 669 CT-----TCAGAAAGGTGTGTACTGGCTGAGATCTGCTCCAGTCCCTCCCTCG 619
Db      :::::
QY 163 laArgPheProProProArgArgCysArgIle--ProAlaGlyProGlyProPro--- 180
Db      :::::
QY 618 GCGCAGCTCCCGAGAACTTTCGAG--ACAGCCTGACGTGCGAGTGGCGGT----- 569
Db      :::::
QY 181 --ArgArgProGlyArgProThrArgGlyThrAla-----ArgTyrArgCysArgGlyC 198
Db      :::::
QY 568 -----CCACGAAAGGCGCAGGTGGAACATGTCGCGATCTTCGCTGTA 523
Db      :::::
QY 198 ysCysAlaProAlaProAlaArgArgSerProAlaIleProAlaAlaArgProAlaCysA 218
Db      :::::
QY 522 GAAACCAAGATGCCCTCAACAGAACAGCTCCGCGAGGTAGA----- 479
Db      :::::
QY 218 laAlaProArgAlaGlyAlaGlyAlaAlaArgProAlaGlyGlnGlySerProAlaArgC 238
Db      :::::
QY 478 -----CCACCGTGTCTCTGTGAACCTTGAGTGTGCACAAATCATCATG---TCGG 430
Db      :::::
QY 238 ysAlaAlaProSerArgProAlaAlaAlaArgSerAlaThrThrGlyArgArgProSerA 258
Db      :::::
QY 429 CACCTCACCGCTTTGCCCTCCACGAGTGTCTTCAAGTCTGTGATCAATCATTAATC 370
Db      :::::
QY 258 laSerArgAlaGlyArgProAlaArgCysProAlaGluArgArgAlaThr----- 274
Db      :::::
QY 369 AAAAGCATCTGAGTGTCAAAATGTACTGCTTTCAAGGCTTGCGCTTCTGCTGC 310
Db      :::::
QY 275 -----AlaProAlaAlaHisArgAlaIleTyrProProProAlaP 287
Db      :::::
QY 309 CGTCAGACCTTTGAGAACTGTCTGCTGCTGAGATGACCACTCCGCTCCGCTGTC 250
Db      :::::
QY 287 roAla-----ProAlaAlaProAlaAlaP 295
Db      :::::
QY 249 CA-----CCTGTTCTGTCTCCAGCACTCATGATCTTTCACACAGGTCGACTTCC 196

```

```

Db      :::::
QY 295 roGlyAlaAlaArgSerArgProAlaGlyPro-----ArgArgGlyArgArgA 311
Db      :::::
QY 195 GCTGGCAGTCCCGCCGCTGACCCCTTACAGAAAGGCGCGGTGGTGGAGAGGTGGCTC 136
Db      :::::
QY 311 rgaArgArgCysArgSerAlaPro-----AlaGlySerGlyGlyAsnPro- 325
Db      :::::
QY 135 CGGCGCGGGGCTCTCGCAGTCTTCGCTCCCGCGCAAGCCATCTCGGCTCCGCTCCGC 76
Db      :::::
QY 326 ArgLysArgArgLysGlnGlyCysCysArgArgArgArgSer-----AlaAlaPro 341
Db      :::::
QY 75 GCATCGGGTCCCGCGCCGCTTCCTCCCGGCGCGCGCGCCGCGCCAGAGT 16
Db      :::::
QY 342 AlAlaAlaProGlyAla-AlaAlaSerProGlySerGlySerProProAlaAlaArgPr 361
Db      :::::
QY 15 CGGAGCGGACCCC 3
Db      :::::
QY 361 ovalAlaAsnPro 365

```

RESULT 8

```

US-09-252-991A-23264
: Sequence 23264: Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252.991A
: PRIOR FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074.788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094.190
: NUMBER OF SEQ. ID NOS: 33142
: SEQ. ID NO 23264
: LENGTH: 568
: TYPE: PRT
: ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23264

```

Alignment Scores:

```

Pred. No.: 5,67e-09 Length: 568
Score: 196.00 Matches: 119
Percent Similarity: 34.73% Conservative: 30
Best Local Similarity: 27.74% Mismatches: 150
Query Match: 6.65% Indels: 130
DB: Gaps: 23

```

US-09-896-522-1 (1-1624) x US-09-252-991A-23264 (1-568)

```

QY 1127 TCAGGAAGCGCTGTCAAGTGTCCAGCAAGTTGAGTGCAGATCTGATTCAC 1068
Db      :::::
QY 9 SerGlySerProAlaThrCysArgSerThr----- 18
Db      :::::
QY 1067 TCCTGATGAGAGAGCGCTGTCTAACACTCCCTGGGGTGGCGCGAGAGAGCACT 1008
Db      :::::
QY 19 -----ThrArgSerThrMetProCysAlaIleTyrProAlaPro--ProAlaThr 34
Db      :::::
QY 1007 GGGTGGGGTGGCGCTCCCGAGGCTCAGTCCCTGAACACACATGCGGGCGGAGACTG 948
Db      :::::
QY 35 GlyArgAlaIlePro-----SerAlaProAspLeuAlaArgArgSer----- 48
Db      :::::
QY 947 CCCTGAGGCTCGGCGCCCTCAGTGGGGTGGCTGC-----TGG----- 908
Db      :::::
QY 49 ProTyrSerValAlaAlaGlyArg-GlyArgCysCysAlaAlaAlaSerSerTyrCysAr 68
Db      :::::
QY 907 -ACTCAAAATGTACCGTTTGCCAGAGTCCAGATCCAGGTGTCCCTGCTCAGAA 849
Db      :::::
QY 68 gthrThrSerArgAlaValProThrArgSerCysAlaIleGly----- 82
Db      :::::
QY 848 AAGTCGCTGTGAGTCCGCGCATTTGAGACCTCTCGGTGCATTGCGATGTACCA 789

```



```

OY 623 GGAGGA----- 629
Db 431 gProglYProProArGProAlaSPProArGlyThrPheProArGlyLeuArGAr 451
OY 630 -----CCTGAGCAGATTCTGAGCAGTACACCACTTCTGTAAGCCGG 673
Db 451 gValGInglYglYAlaProGlYAlaGlYLeuArGProArGhIbHs----- 466
OY 674 CTTTCGAGAGATTCTGCTCCGACAAGATAGCCATGTATGATCCACGAGAG 733
Db 467 -----HisGlYglYAlaProAlaSPLeuArG-----ProHisArGglYArGl 481
OY 734 TGGACATAT-----GTTGCCATCAACCTGATGTCGACGACATTC 784
Db 481 YgluArGyAlaProProArGglYValHisArGlyHisArGAlaAlaArGArGSPArP 501
OY 785 TGAATGGAGCATC-----TGCMAATGGCACCGAGAGAGTCCCATG 826
Db 501 rOlEuGlYArSPLeuArGProGlYglYArSPArGAlaValHisIbArGlInglYgl 521
OY 827 GCGGAGCTAC----- 837
Db 521 YglYglYleuArGHisAlaProArGAlaSerGlYleuArGThrGlYAlaIaIeUglY 541
OY 838 -----AACGACCTTTCTGAGCCAGGAGCACACCTG 871
Db 541 lAaRgAlaProArGAlaGlYglYArGArGArProArGArGSPProAlaHisArGAlaG 561
OY 872 GGATGCTGACCTTCGCAAGCGTCACTTTGAGTCCAGACAGACCCCACTGAGGG 931
Db 561 lYglYleuAlaSPglYlnArGValHis-----SerGlInProAlnArGlY 577
OY 932 CTGCGAGCTCAGGCGTCTCCCGCCGACATGTGTTCAGGAGCTGAGG 991
Db 577 rGlEuArGHisAlaGlYProAlaGlYAlaGlYHis-----GlInProArSP- 591
OY 992 AGCGCCACCAACCCACTGCTCTCTGCGCGCACCCAGGAGAGTGTTCAGCAGG 1051
Db 592 -----ProValGlYArSPAlaAlaSPProGlYglYlnGlY 604
OY 1052 CCTTCCTCACTCAGAGTGAAGAACTCAGATGTGTCACTCACTCACTCTGCGGAC 1111
Db 604 lA-----AspGlYglYlnAlaProGlYSPAr 612
OY 1112 TGACAGCGCT-----CCTGAGGTTTTCAGCC 1138
Db 612 rGlAlaGlYleuArGProHisArGSPArGArGArGlYleuArGProArSPAlaIaGlY 632
OY 1139 ACTTAGGCTCTGCGGTTTAAAGATCCCTTAGTCACTGAGAATGCCACAGATGTG 1198
Db 632 rGlEuArGglYlnLeuArGSPAlaGlYlnAlaIaGlYProValAlaGlYProGlYlnGlY 652
OY 1199 CAGGAGCGTGGAGGC 1215
Db 652 lAglYglYAlaIaIaGlY 657

RESULT 10
US-09-252-991A-32424
; Sequence 32424, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32424

```

```

; LENGTH: 341
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32424

Alignment Scores:
Pred. No.: 2,09e-08 Length: 341
Score: 188.50 Matches: 103
Percent Similarity: 31.62% Conservative: 26
Best Local Similarity: 25.25% Mismatches: 139
Query Match: 6.40% Indels: 140
DB: Gaps: 25

US-09-896-522-1 (1-1624) x US-09-252-991A-32424 (1-341)
OY 1022 CCGAGAGGA-----AGCATGGGTGGTGGGCGTCCCGCAGGCTCAGTCCGTAAC 969
Db 2 ProArGSPAlaCySPThSerAlaCySPArGTrProProGlYSerAlaIa----- 19
OY 968 ACATGCCGCGGAGAGACCTGCTGAGGCTCGGACGCCCTCAGTGGGGTGTGCTGTG 909
Db 20 -----MetProGlYglY-----AlaPro----- 25
OY 908 GACTCCAAATGTGACCGTTTCCAGAGTCAATC-----CCAGGGTGTCCCTGCTCA 852
Db 26 ---ThrAlaIeSerYSPAlaSerAlaAlaIleSerProSerThSerProGlYAla 44
OY 851 GAAAGGTCCGCTTGATGCTCCGCGCCATCGGACCTCTCGGTCGATTTGCAGATGTCA 792
Db 45 ProArSP- ProProThSerThSerCysPheAlaArGThSerAlaIaCys----- 61
OY 791 CCATTCAGATATCTCTGATGTCTGCACAGATCAGTGTGATG-----CAACCATATTGTC 735
Db 62 -HisTrpThrCySPTrArGProAlaProAlaSerProAlaTrIleProProIYrCys-- 80
OY 734 ACTTCGTGGGATGATACATCGATCGATCTTTGTGGCAGGACGAATCTTCGAAG 675
Db 81 -----LeuProArGSPProAlaY 87
OY 674 GCGGCGTTCACGAAGGTGTACTGCGTCAAGATCTGCTCCAGTCCCTCGGCGC 615
Db 87 SPProArGlySerAlaTrIYrCySPArGArSer-----ProProArGSPAr 102
OY 614 ACGT-----CCGAGAACTCTTCGAGACGCTGACGTGAGTGTGCTCC 567
Db 102 eArGProThTrIYrProGlYSerCys-----AsnArGSPAr 114
OY 566 ACGAAGCGCGCAGGTGAATGTCC-----GGATCTCTGCTGTGAACA-- 518
Db 115 ---ThrGlYSerAlaThrValAlaProAlaGlYThrSerSerProArGSPArGlA 133
OY 517 -----CAAAGTCCCTCAACAGAACCA-----CGTCCGAGGG 483
Db 133 rArGThrSerAlaTrIYrThrArGSPArGSPArGSPArSerAlaThrArGAlaProP 153
OY 482 TAGACCAACCGTGG-----TCTCTGTGAACCTTGAAGTGTGACA 444
Db 153 rSerProGlYTrIYrCySPThrThrCySPglYThrAlaSerProAlaArGHisSerTr 173
OY 443 AAATCATAGTGGCGACCTTCACACGTTTGGCCCTCAGCATGTTCTCAAGTCTGTGC 384
Db 173 rAlaArGAlaAlaIaIaThrProThrPhe-----ProGlYCySPThrArGSPAr 191
OY 383 ATCAATCATTTCAAGGATGTGATGTCAAAATTGATCGTCTTCAAG----- 329
Db 191 rAlaArGSPArSPArSPArSPArSPArSPArSPArSPArSPArSPArSPArSPAr 210
OY 328 ---CCTTGGCTTCTGCTGCTGCGGTCAGAGCTTGTAAACCTGCTGCTGCTGAGATG 273
Db 210 YrPheProTrP-----CysArGSPArAlaPro----- 218
OY 272 ACCACCTTCGCGTCCGCTGTTCACCTGT----- 242

```

```

Db      219 -ProArgAlaGlyAlaThrSerProProArgTrpProThrTrpCysSerArgArgProAr
QY      241 -----TCTGTCCAGCAACCTCCAGTCTTCTCA 213
Db      238 gthrArGArgLeuProTrpAsnCysAlaAlaThrPheProAlaAlaProsmnLatrrva 258
QY      212 CACAGGTGCACATCCCGCTGAGCATGCGCCGCTCCATCAGAGAGGCGCGCTGG 153
Db      258 lAlaArgArgAlaThrCysAlaCysCysArgAlaSerAlaCysThrTrpArgArgAlaAl 278
QY      152 TGGGACGAGTGGCGCTCCGCGCGGCGCTCTCCGAGTCTTCGCTCCCGCGAAGCATC 93
Db      278 aArg-----ArgArgGlyLysSerSerCysArg-----Proth 289
QY      92 TGGGCTCCGCTCCCGGAGTCCGCGCGCGCGCG-----56
Db      289 rArgAlaTrpThrAlaArgAlaSerSerProArgSerGlyArgCysSerAlaSerThrse 309
QY      55 -----CCCTTCCCGGCGCGCGCGCGCGCGCGCGCGCG 24
Db      309 rAlaThrGlyArgTrpThrGlyArgPheProArgProThrCysArgArgArgTrpProth 329
QY      23 GCCGAGGTCCGAGGCGACCCCA 2
Db      329 rProAlaAlaAlaProThrPro 336

RESULT 11
US-09-252-991A-31128
/ Sequence 31128, Application US/09252991A
/ Patent No. 6551795
/ GENERAL INFORMATION:
/ APPLICANT: Marc J. Rubenfield et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
/ FILE REFERENCE: 107196.136
/ CURRENT APPLICATION NUMBER: US/09/252,991A
/ PRIOR FILING DATE: 1999-02-18
/ PRIOR APPLICATION NUMBER: US 60/074,788
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 60/094,190
/ NUMBER OF SEQ ID NOS: 33142
/ SEQ ID NO 31128
/ LENGTH: 375
/ TYPE: PRt
/ ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31128

Alignment Scores:
Pred. No.: 2 41e-08 Length: 375
Score: 188.00 Matches: 112
Percent Similarity: 33.49% Conservative: 33
Best Local Similarity: 25.87% Mismatches: 133
Query Match: 6.34% Indels: 155
DB: 4 Gaps: 19
US-09-896-522-1 (1-1624) x US-09-252-991A-31128 (1-375)

QY      9 GCCTCCGACCTCGCGCTGCGCG-----GCCGCG 38
Db      27 AlAlaAspProGlyAlaGlyArgArgGlySerProGlyAlaCysProGlyAspArgArg 46
QY      39 CGGCGCGGGAAGGGCGGCGCGCGGACCCGATGCGCGGAGCGGA-----86
Db      47 GluLeuGluGlyCysAlaLeuAlaSerGluProGlyProGlyAlaGlyCysAlaAlaVal 66
QY      87 -----GCCCGAGATGCTTGGCGGAGGAGAACTGCGAGAGCCC 128
Db      67 ArgSerArgLeuArgThrArgArgGlyAlaGlyGlyArgArgArgGlnArgPro 86
QY      129 CGCGCGCGGA-----GCCCGA-----CCGTCGACCGACGCGCGCTT 164
Db      87 AlAlaGlyPheProArgGlyGlnPheProArgProTrpArgSerLeuSerGlyLeu 106

```

```

QY      165 CTTGATAGGGGTAGCGGCGGACCTGCAGCGGGAAGTCGACCGGTGTGAGAAATCAT 224
Db      107 TrpArgArgGly-----110
QY      225 GAGTGTCTGGGACAGAACAGAGGTGAGACGCGGACGGAAAGTGTATCTTCAGCCA 284
Db      111 -----AlaAspAlaThrArgGlyProGlyGlyLeuProArgGlyThnHisProGlyGly 128
QY      285 GGCACAGTTCTA-----CAAGTCTGACGGCACA 314
Db      129 GlyGlnArgLeuLeuHisArgProAlaGlyGlyLeuPheArgProGlyArgProGlyArg 148
QY      315 GCAGAGGCGCAAGGCTTGAAGAGACATGATTTTACATCCAGATGCTTTATTA 374
Db      149 ValGluGlyArgAla-----153
QY      375 TGATTTGATGCACAGGACTCTGAGAACATCGTGAGGCGCAAAAC-----GGTGAGGTGCC 431
Db      154 -----AspArgAlaAspHisArgGlyGlyAlaHisValArgAlaGlyLeu 168
QY      432 GACCTTGAATTTGTGTACACATCAAGTTTACAGAGACACAGGTGTCTACCTTCGGA 491
Db      169 Asp-----ProValGly 172
QY      492 CGTGTCTGTGTGAGGCGATCTGTGTCTTACAGCCAGAGATCGGCGACATGTTCCA 551
Db      173 GlnGlySer-----ArgArgAlaGlyAsp-----180
QY      552 CTTGCCCTCTTGTGAGACGACGATCCGACGTCAGCTGTCTCGAAGATTTCTCCGGA 611
Db      181 -----ArgPheHisGlnValArgArgGlyAlaArgValProProArgAlaGly 196
QY      612 CGTGGCGCG-----AGGAGGACCTGAGACAGATTCTGACGACATAC 656
Db      197 ArgArgGlnGlyGlyGlyAspArgArgAlaGlyProGlyProGlySerArgAlaLeuHis 216
QY      657 -----CACCTTGTGAAAGCGCGGCTTCGAGAGATTCGCTGCC 695
Db      217 ArgGlyGlnGlyArgGlnCysHisProArgArgGlyLeu-----ArgProAla 232
QY      696 GACAAAGATATGCCGATGT---GATCATCCACAGAGAGTGACAAATATGTTGCCAT 752
Db      233 AspAspProAlaArgArgCysLeuArgHisProArgGlnAlaGlyAlaValTrpLeuGln 252
QY      753 CAACCTGATGTGCACACATCCAGGACATTTGAAATGTGACATCTGCAAATGSCACCG 812
Db      253 ArgArgGlnArgValGlyValProGly-----Leu 262
QY      813 AGGAGGTCCTCAATGGCGGAGC-----TACAAAGC 842
Db      263 ArgArgValGlnGlnAla-ProAlaValLeuProProLeuProAlaSerGlyPheHisArgse 282
QY      843 GACCTTTTTCGAGCCAGGGAACCACTCGGATGTGACCTCTGCAAAACGGTCACATTT 902
Db      282 rSerArgAspGlyProGlyThrGlnArgArgValGlyGlnGlnGlyProArgAlaHisArg 302
QY      903 GGAGTCACAGACAGACCCCACTGAGGCGCTCGGAGGCTCCAGGCGAGGTCTCCCGCG 962
Db      302 gAlaThrAspProArg-----SerProAlaGln-AlaGlyGlyArgProG 317
QY      963 GCATGTGTGTTCAGGACTGAGCCTTGAGACGCCCAACCCACACCCACTGCTCTCCGG 1022
Db      317 LylLeuArgValArgProGlyGlyArgGlyAlaProLeuHisGlyAsn---LeuSerGln 336
QY      1023 CGCACCCCGAGGAGGTATAGACAGGAGCTTCTCCATCTCAGAGAGTGAAGAACTCAGATG 1082
Db      336 lArgProGlyAlaAspProArgArgArgLeuArgThrleArgSerGlnAlaArgP 356
QY      1083 TGTCACTCACTCAACTGCTGGACACTGAC 1115
Db      356 roSerSerArg-AlaPheSerTrpGlyThrGln 366

```


QY 397 -----TCAGAGCTCTGCATCAATATATCAAGGACCTGGATGTCAAAATTGT 344
Db 111 ArgProGlyArgArgCysAlaSerArgArgGlyArgArg----- 123
QY 343 ACTGCTCTTCAAGGCTTGG-----CCTTGTGCTGCCGTGAGACCTTTGAGAAC 290
Db 124 -----AlaAlaArgProGlyCysAlaProAlaAlaGlyLeuSer---ProArgHisPro 140
QY 289 TGTCTGCTGAGTGAAGTGAACACCTTCGCTGCC-----GCTGTTCGA---CCTGTCTT 239
Db 141 AlaProArgArg-----ProProAlaArgAlaArgGlnAlaProProAlaProArgArg 158
QY 238 GTCCAGAGAACTCCATGATCTTCTTCACACAGGTCGACTTCCGCTGCACTGCCAGTCCGCC 179
Db 159 ArgProAlaValProGlySer-----TTPArgCysArgAla 170
QY 178 TCACCCCTATCAGAGAGGCGCTGTGTGCGAGCGGTGCGCTCCGCGCGCGCTGCTGC 119
Db 171 GlyProVal-----ArgAlaAlaProAla-----ProGlyArgGlyTyrArg 184
QY 118 AG-TCTTCGCTCCCGCGAAGCATCTCGGCTCCGCTCCGCGCATCGGATCCCGCG 60
Db 185 ArgArgSerProProAlaAlaAlaThrAlaProAlaAlaGlyAlaGlyArgThrSerAla 204
QY 59 ----CCCCGCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 12
Db 205 GlyAlaAlaProAlaGlyArgArgProAlaAlaProProAlaAlaAlaProGly 221

RESULT 15

US-09-252-991A-31760
Sequence 31760 Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 31760
LENGTH: 1706
TYPE: PR
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-31760

Alignment Scores:

Pred. No.: 9.63e-08 Length: 1706
Score: 184.50 Matches: 129
Percent Similarity: 32.82% Conservative: 41
Best Local Similarity: 24.90% Mismatches: 185
Query Match: 6.26% Indels: 163
DB: 4 Gaps: 25

US-09-896-522-1 (1-1624) x US-09-252-991A-31760 (1-1706)

QY 1388 AAGTACTTGTTCACAAACCTTTGAGTTATGTGTGAGTGTGTAAGAACAGATC--- 1332
Db 1230 LysAlaValArg1LevalThrArgGlnGlnAlaLeaThrCysGluArgAspLeuGln 1249
QY 1331 AGACTGAAAAAACTCTCTCCACTGTGGTTCACCTGTCAACAAACATCAGGCGACCA 1272
Db 1250 ArgLeuAspArgAlaLeuAlaArgTyr-----GluArgGlnAlaSerArg 1264
QY 1271 GTGTCTAGGCTGTCTCTCAATTTCCCAATATATGTCCTCACAATCCAGAACCT 1212
Db 1265 LeuAlaGlnLeuSerAsp-----AlaGlnArgAlaAlaAlaLeaAla 1278
QY 1211 CCAGAGCTTCTGTGACATTTCTGTGTCATTCTCAGTACCTAAGAGGATCTTTAAACCGC 1152

Db 1279 ArgArgAlaSerLeuHisAlaLeuAlaArgThrGlyThrLeuAlaGlyAlaThrThr 1298
QY 1151 AACGAGCTTAAGTGGCTGAAAACCTCAGAACGCTGTGCTAGTCCACAGATGTGATC 1092
Db 1299 GlyGlnAspArgArgArg-1LeProGlnAlaArgProArgArgThrArgArgAlaArg-- 1317
QY 1091 TGAATGACAC-----ATCTGAGTTCCACTCTGAGTGAAGAAAGGCTCG 1047
Db 1318 -----HisProProGlyArgGlyLeuThrProAlaThrProProProGlySer 1335
QY 1046 CTGTAACACTCCCTCGGGGTGGCGAGAGAAAGACAGTGGTGTGGTGGGCTCCCA 987
Db 1335 GglnArgProAlaProGlyProArgArgGlnAlaArgCysArgGlnArgGlySerProAl 1355
QY 986 GGCTCAGTCCCTGAACACACATCCGCGCGGAGACCTGCTGAGGCTGCG----- 935
Db 1355 aaenProAla-----SerAlaGlyArgArgArgProProArgArgArgGlySerPr 1372
QY 934 -----CAGCCCTCAGTGGGCTGTGCTGGAC 906
Db 1372 oAlaArgProArgPheCysArgProArgGlnArgProArgArgGly-----Th 1388
QY 905 TCCAAATGTACCGCTTTGCCAGAGTTCAGCATCCAGGAGTGTCCCTGCTCAGAAA 846
Db 1388 rPro-----GlnArgSerProAlaArgThrArgProAlaProGlnAspAr 1403
QY 845 GTCCGCTTGAGTCCGCCCATGAGACCCCTCGGAGCATTTGACAGATGTCAGATTC 786
Db 1403 gArg-----AsnAlaHisHis-- 1408
QY 785 AGAATGCTGTGATGTGCTGACAGATCAGTGTGATGAGCAACATATGTCCACTCTCGT 726
Db 1409 -----ProGlyThrValAlaArgProPro----- 1416
QY 725 GGGATATACATCGGCATTACTTCTTGTGCGAGGACAGACTCTCGA----- 677
Db 1417 -----AlaThrGlyArgProAlaArgThrAlaProG1 1427
QY 676 -----AGCGCGCTTACGAAAGTGGTGTACTGCGTCAAGTATGCTCCAGGCTC 624
Db 1427 yAlaHisArgProAlaHisArgArgThrAlaThrAlaAlaArgArg----- 1442
QY 623 CCTCGGCGCAGTCCCGGAAACTTTGAGACAGCCTGAGTGGAGTGGTGTCCAG 564
Db 1443 ----GlyGlnArgProSerValProArgAlaThrGlyThrArgArgSerArg----- 1458
QY 563 AAGAGCGCAGGTGAAACATGTCCCGGATCTCTGCTGTGAACACCAAGATGCCCTCA 504
Db 1459 -----ThrAlaProGlyAlaPro-----GlnProAlaAlaArgG1 1470
QY 503 AACGAAACACGTCGCGCAG--GGTAGACACCGGTGCTGTGTGAACCTTGAGTGTGC 447
Db 1470 nProGlyProArgProArgProGlyArgPro----- 1480
QY 446 ACAAAATCATAGTGGCAGCTTCACCGCTTTCCTCCACGATGTTCTTCAGAGTCTG 387
Db 1481 -ArgThrProAlaAlaThrProAlaProGlyThrProAlaArgProArgG1ArgSerAr 1500
QY 386 TGCATCAATCAT-----TATCAAGGATCTGATGTGTA 351
Db 1500 gArgProArgArgArgArgThrArgArgThrAlaProAlaArgArgArgLeuProAlaAr 1520
QY 350 AAATGTACTGTCCTTCAAGGCTTGGCTGCTGCTGCGCTGAGACCTTGTGAAC 291
Db 1520 gGlnArgSerAlaAlaThrArgArgThrHisArgAlaLeuGlnArgHisProHisArgPr 1540
QY 290 CTGTCTGCTGAGTGAAGCAACCTTCGCTGCGCTGCTGTTCACCTCTGTCTGTCCAGC 231
Db 1540 oProProAlaThrGlyArgPro-----ArgProProProGlyArgAlaProG1 1556
QY 230 AACTCATATCTTTCACACACAGGTGACTTCGCGCTGAGAGTGGCGCGCTCACCCCT 171

```
Db      1556 Y---ProArgGlnProArgLeuArgSerProArgArg----- 1567
Qy      170 ATCAGGAAGGCCCGCTGTGCGGAC----- 146
Db      1568 -HISGlyHisArgLeuGlyAlaAspArgGlnGlyArgProAlaGlnThrArgHisProAr 1587
Qy      145 -----GGTCGGCCTCCGGCGCGGGGC---TCTCG 120
Db      1587 gLeuArgProGlySerArgArgGlnGlyArgGlnArgProProAlaThrAlaArgArgAr 1607
Qy      119 CAGTCTTGCCCTCCCGCGAGCCATCTGAGCTCCGCTCCCGCGCATCGAGTCCCGCG 60
Db      1607 gProGlnArgGlnProArgGlnProAlaArgPro-ArgHisArg-----AsnProLeuV 1625
Qy      59 CCGGCCCTTCCCGGGCCCGGGCGCGCCGCCCGCCAGCCCGAGTCCGAGG 10
Db      1625 a1ArgArgValProThrProProGlyProAlaArgArgProGlyArgArg 1641
```

Search completed: November 25, 2003, 07:42:32
Job time : 70.812 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - protein search, using frame_plus_n2p model

Run on: November 25, 2003, 02:31:28 ; Search time 64.4182 Seconds

(without alignments)
8003.075 Million cell updates/sec

Title: US-09-896-522-1

Perfect score: 2967
Sequence: 1 gtcgggctgcctccgaccc.....ccagctctgtcggggccag 1624

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame_n2p.model -DEV=x1h
-O=/cgn2.1/USPTO.spool/US09896522/runat_21112003.184103.2805/app_query.fasta_1.2830
-DB=A.Geneseq.19jun03 -QFMT=fastan -SUFFIX=n2p.rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=bloms62 -TRANS=humand0.cdi
-LIST=45 -DOCALIGN=200 -NOR=score-pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09896522@cgn1.1.112.@runat_21112003.184103.2805 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONCLOG
-DEV TIMEOUT=120 -WARM TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

A.Geneseq.19jun03: *
1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT: *
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT: *
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT: *
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT: *
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT: *
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT: *
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT: *
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT: *
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT: *
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT: *
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT: *
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT: *
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT: *
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT: *
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT: *
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT: *
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT: *
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT: *
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT: *
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT: *
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT: *
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT: *
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT: *
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1551	52.3	296	22	AA1980
2	1450	48.9	277	22	AA1981
3	1450	48.9	277	22	AA1982
4	1450	48.9	277	22	AA1983
5	1372.5	46.3	276	22	AA1984
6	1363	45.9	260	22	AA1985
7	1325	44.7	277	22	AA1986
8	951	32.1	190	23	AA1987
9	934	31.5	261	22	AA1988
10	934	31.5	261	22	AA1989
11	930	31.3	337	21	AA1990
12	925	31.2	337	21	AA1991
13	914	30.8	335	22	AA1992
14	738	24.9	260	22	AA1993
15	610.5	20.6	326	22	AA1994
16	581	19.6	481	21	AA1995
17	581	19.6	490	21	AA1996
18	581	19.6	512	21	AA1997
19	579	19.5	120	23	AA1998
20	572	19.3	500	21	AA1999
21	572	19.3	511	21	AA2000
22	554.5	18.7	578	22	AA2001
23	551.5	18.6	548	22	AA2002
24	548	18.5	614	22	AA2003
25	548	18.5	614	22	AA2004
26	546.5	18.4	548	21	AA2005
27	535	18.0	540	21	AA2006
28	529	17.8	415	21	AA2007
29	529	17.8	433	21	AA2008
30	529	17.8	441	21	AA2009
31	529	17.8	443	21	AA2010
32	529	17.8	461	21	AA2011
33	529	17.8	469	21	AA2012
34	518	17.5	477	21	AA2013
35	518	17.5	500	21	AA2014
36	517.5	17.4	466	21	AA2015
37	507.5	17.1	476	21	AA2016
38	507.5	17.1	498	21	AA2017
39	507.5	17.1	499	21	AA2018
40	507.5	17.1	521	21	AA2019
41	507	17.1	465	21	AA2020
42	507	17.1	487	21	AA2021
43	490.5	16.5	274	21	AA2022
44	490.5	16.5	292	21	AA2023
45	490.5	16.5	300	21	AA2024

ALIGNMENTS

RESULT 1
ID AA1980 standard; Protein; 296 AA.
XX AA1981
AC AA1982
XX AA1983
XX AA1984 (first entry)
DT 22-OCT-2001
XX Human polypeptide SFG ID NO 6219.
XX Human, noctropic; immunosuppressant; cytostatic; gene therapy; cancer;
DE peripheral nervous system; neuropathy; central nervous system; CNS;
XX Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
OS Homo sapiens.

[illegible]

Db	61	LysIleMeGluIleuLeuEnglyGlnSngIuValIGluGlnArgGlnIaArgIyValValIle	80
Qy	277	CTGAGCCAGGACAGGTTCTTACAAAGGCTCTGACGGCGAGACAGAAAGGCCACCGCTTGAA	33
Db	81	LeuSerGlnAspIrrpRheTyrlYsValIleuThrIaGluGlnLysAlaLysAlaLeuIys	10
Qy	337	GGACATPACAAATTTTACCATCCAGATGGCTTGATTAATGATTTGTGACAGCACTCG	39
Db	101	GIYGlntTyAsnAspRheAspHisProAspAlaRheAspAsnAspLeuMetHisrGlnLeu	12
Qy	397	AAACACATCGTGGAGGGCAAAAACGGTGAAGTGCCGACCTATGATTTTGTGACACACTCA	45
Db	121	LysAsnIleValIGluIuIystrThrValIGluValProThrTyAspPheValThrHisSer	14
Qy	457	AGGTTACACAGACACACGGTGGCTTACCTCGGAGAGCTGGTCTGTGTAAGGCATCTTG	51
Db	141	ArgIeuProGluThrThrValValTyProAlaAspValIleuPheGluIuIyIleLeu	16
Qy	517	GTTGTTCCACGACGAGAGATCCGGGACATGTTCCACTGCGGCTCTGCTGGACACCGAC	57
Db	161	ValPheTySerGlnGlnIleArgAspMetPneHisLeuArgIeuPheValAspThrAsp	18
Qy	577	TCCGACGCTCAGGCTGTCTCGAAGAGTTCCTCCGGGACGTGCGCGGAGGAGGACCTGAG	63
Db	181	SerAspValIaArgIeuSerArgValIleuArgAspValIaArgArgIyArgAspLeuGlu	20
Qy	637	CAGATTCTGACGCAATCACACACTTTCGTGAACCCGGCTTGAGAGAGTTCTCCCTCGCG	69
Db	201	GlnIleuThrGlnTyrlThrThrPheValIyAspProAlaPheGluGluIuPheCysIeuPro	22
Qy	697	ACAAAGAGATCCCGATGTGATCATCCACGAGGAGTGAGGACAAATPAGTGGTCATCAAC	75
Db	221	ThrIySyrTyrlaAspValIleIleProArgIyValaAspAsnMetValaIaIleAsn	24
Qy	757	CTGATCGTCGACGACATCCAGACATTTGTAATGTGATCTGCAATATGGACCCGAGGA	81
Db	241	LeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysIystrIrrHisArgGly	26
Qy	817	GGGTCCAATGGGCGAGACTACAAAGCGAGCTTTTCTGAGCCAGGGGACCACTCGGGATG	87
Db	261	GlySerAsnGlyArgSerTyrlYsArgThrPheSerGluProGlyAspHisProGlyMet	28
Qy	877	CTGACCTTCGGCAAAACGCTACACTTTGGAGTCCAGAGACAGCCCGAC	94
Db	281	LeuThrSerGlyIyAspSerHisLeuGluSerSerArgProHis	296
RESULT 2			
AAM39502			
ID	AAM39502 standard; Protein; 277 AA.		
XX	AAM39502;		
AC	22-OCT-2001 (first entry)		
DT	Human polypeptide SEQ ID NO 2647.		
XX	Human, neotropic; immunosuppressant; cytostatic; gene therapy; cancer;		
KW	peripheral nervous system; neuropathy; central nervous system; CNS;		
KW	Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;		
KW	amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemocarcin;		
KW	chemokineic; thrombolytic; drug screening; arthritis; inflammation;		
XX	Leukaemia.		
XX	Homo sapiens.		
OS	MO20015312-Al.		
XX	26-JUL-2001.		
XX	26-DEC-2000; 2000MO-US34263.		
XX	21-JAN-2000; 2000US-0488725.		
PR	25-APR-2000; 2000US-0552317.		

PR 09-JUL-2000; 2000US-0598042.
 PR 19-JUL-2000; 2000US-0620312.
 PR 03-AUG-2000; 2000US-0653450.
 PR 14-SEP-2000; 2000US-0662191.
 PR 19-OCT-2000; 2000US-0693036.
 PR 29-NOV-2000; 2000US-0727344.
 XX
 XX (HYSE-) HYSEQ INC.
 XX
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
 PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J;
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;
 XX
 DR WPI: 2001-442253/47.
 DR N-PSDB; AA158658.
 XX
 PT Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 PS Example 4, SEQ ID NO 2647; 10078bp; English.
 XX
 XX The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AAM38642-AAM42213) with neurotropic,
 CC immunosuppressant and cyostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: Immune system suppression,
 CC Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemia and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 CC
 XX SQ Sequence 277 AA;
 Alignment Scores:
 Pred. No.: 4.1e-110 Length: 277
 Score: 1450.00 Matches: 277
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 48.87% Indels: 0
 DB: 22 Gaps: 0
 US-09-896-522-1 (1-1624) x AAM39502 (1-277)

QY 454 TCAAGTTACAGAGACCAAGGTGTTACCTGCGAGCGTGTCTGTTGAGGCGATC 513
 DB 121 SerArgLeuProGlnThrValValTyrProAlaAspValValLeuPheGluGlyLe 140
 QY 514 TTGTGTTCTACAGCAGAGATCCGGACATGTTCACTTCGAGCTTCTGTCGACAC 573
 DB 141 LeuValPheTyrSerGlnGlnIleArgAspMetPheHisLeuArgLeuPheValAspThr 160
 QY 574 GACTCCGACGTACAGCTGTCTTGAAGAAGTCTCCGGAGACGTGCGGAGGAGGACCTG 633
 DB 161 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 180
 QY 634 GAGCAGATTCTGACCCAGTACACCACTTCGGAAGCGCGCTTGGAGGAGTTCGCGC 693
 DB 181 GluGlnIleLeuThrGlnThrThrPheValTyrProAlaPheGluGlnPheCysLeu 200
 QY 694 CCGACAAAGAAAGTATCCGATGTGATCATCCACAGAGAGTGAGCAATATGCTGGCATC 753
 DB 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
 QY 754 AACCTGATCTGCAGACATCCAGACATCTGAAATGCTGACATCTGCAATGCGACCGA 813
 DB 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlyAspIleCysLysThrPheArg 240
 QY 814 GGAGGGTCCAAATGGCGGAGCTACAAAGCGACCTTTTCGACCCAGGGGACACCCCTGG 873
 DB 241 GlyGlySerAsnGlyArgSerTyrLysArgThrPheSerGlnProGlyAspHisProGly 260
 QY 874 ATGCTGACCTCTGGCAAAAGGTACATTTGAGTGCAGACGAGACCCGAC 924
 DB 261 MetLeuThrSerGlyLysArgSerHisLeuGlnSerSerArgProHis 277
 XX
 XX RESULT 3
 XX AAB93941
 XX ID AAB93941 standard; Protein; 277 AA.
 XX
 XX AC AAB93941;
 XX
 XX DT 26-JUN-2001 (first entry)
 XX
 XX DE Human protein sequence SEQ ID NO:13952.
 XX
 XX KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.
 XX
 XX OS Homo sapiens.
 XX
 XX PN EPI074617-A2.
 XX
 XX PD 07-FEB-2001.
 XX
 XX PF 28-JUL-2000; 2000EP-0116126.
 XX
 XX PR 29-JUL-1999; 99JP-0248036.
 XX PR 27-AUG-1999; 99JP-0300253.
 XX PR 11-JAN-2000; 2000JP-0118776.
 XX PR 02-MAY-2000; 2000JP-0183767.
 XX PR 09-JUN-2000; 2000JP-0241899.
 XX
 XX PA (HELI-) HELIX RES INST.
 XX
 XX PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
 XX
 XX DR WPI: 2001-318749/34.
 XX
 XX PT Primer sets for synthesizing polynucleotides, particularly the 5602
 PT full-length cDNAs defined in the specification, and for the detection
 PT and/or diagnosis of the abnormality of the proteins encoded by the
 PT full-length cDNAs -
 XX
 XX PS Claim 8, SEQ ID 13952; 2537bp + CD ROM; English.
 XX

FT	Modified-site	247..249	/label= Protein_kinase-C_phosphorylation_site
FT	Modified-site	251..254	/label= Casein_kinase_II_phosphorylation_site
FT	Modified-site	260..265	/label= N-myristoylation_site
FT	Modified-site	264..266	/label= Protein_kinase-C_phosphorylation_site
FT	Modified-site	264..267	/label= Amlaction_site
FT	Modified-site	268..271	/label= Casein_kinase_II_phosphorylation_site
FT	Modified-site	273..275	/label= Protein_kinase-C_phosphorylation_site
XX			
PN	WO200202761-A2.		
PD	10-JAN-2002.		
PF	28-JUN-2001; 2001WO-US21063.		
PR	30-JUN-2000; 2000US-216503P.		
PA	(MILL-) MILLENNIUM PHARM INC.		
PI	Glucksmann MA;		
PI	WPI; 2002-140091/18.		
DR	N-PSDB; AAD27186.		
PT	New isolated human uridine kinase family polypeptide 57658, useful for treating hematopoietic neoplastic disorders and disorders of neurons, heart and blood vessels		
PT	heart and blood vessels		
XX			
PS	Claim 4; Fig 1a; 103pp; English.		
CC	The patent discloses human uridine kinase-like polypeptides, designated 57658 and polynucleotides encoding such proteins. 57658 DNA and proteins are useful for developing diagnostic and therapeutic agents for 57658-mediated or related disorders such as haematopoietic neoplastic disorders (e.g. leukaemia), haematopoietic disorders (e.g. psoriasis, dermatitis, diabetes mellitus, rheumatoid arthritis, disorders of bone (e.g. rickets, osteoporosis), heart (e.g. sarcoma, myocardial infarction), blood vessels (e.g. hypertension, atherosclerosis) and neurons (e.g. Huntington's disease, Alzheimer's disease, Parkinson's disease). 57658 sequences or their antibodies are useful in screening assays, detection assays (e.g. forensic biology) and predictive medicine (e.g. diagnostic assays, prognostic assays, and monitoring clinical trials and pharmacogenomics). They are useful as reagents for diagnosing and treating 57658-mediated disorders. 57658 DNAs are useful in gene therapy, in chromosome mapping, to identify an individual from a minute biological sample (tissue typing) and to aid in forensic identification of the biological sample. The present sequence is human 57658 protein.		
XX			
XX	Sequence	277 AA;	
Alignment Scores:			
Pred. No.:	4,1e-110	Length:	277
Score:	1450.00	Matches:	277
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	48..87%	Indels:	0
DB:	23	Gaps:	0
US-09-896-522-1 (1-1624) x AAL16592 (1-277)			
QY	ATGGCTTCGGCGGAGCGGAGAGACTGCGAGAGCCCCCGGCGCGGAGCGGACGTCGCGAC		
Db	1 MetAlaSerAlaGlyGlyGlyAlaPyrGcgGlnSerProAlaProGlyAlaAlaPyrProHis		
QY	154 CAGCGGCGCTTCTCTGATAGGGGCGAGCGCGGCGACATGCGAGCGGAGGAATCGACCTGTGT		
Db	21 GlnArgProHeuLeuIleGlyValaSerGlyGlyTrpAlaSerGlyLysSerThrValCys		

OY	214	GAAAGATCATGAGTTGCTGGGAGAGAAACGAGGTGGAAACGGGGACCGGAGGTGTC	273
Db	41	GIULYAILLENEGCIULEUENUGLYGIMBNGLUVALGIUGINATGGLNAGLYSVAL	60
OY	274	ATCTTAGCCAGGACGAGTTCTTACAAGGCTCCGTGACGGCAGACGAGAAGGCCAAGCCCTTG	333
Db	61	IIIEUSeRGlnsPhRGheTYrtyrValleuthrAlaGluGlnlySAlaLySAlaLeu	80
OY	334	AAAGACAGTACATTTTGGACCATCCAGATGCTTTGATTAATGATTTTGATGACAGCACT	393
Db	81	LYGGLYGLINTYASnheAspHisProAspAlaPheAspAsnMetHisArgThr	100
OY	394	CTGAAGAACCTGGTGGAGGGCAAAACGGTGGAGGGCCGACCTATGATTTTGTGACACAC	453
Db	101	LEULYASnIleValGIULGlnLYSThrValGIULValProMTrTYrAspPheValThnHis	120
OY	454	TCAAGGTACACAGACACACAGGTGGTCTTACCCCTGGACGTCGTCTTCTGTTTGAAGGCATC	513
Db	121	SeArRGleuProGlnThrValValLYrProAlaAspValValleuPheGluGlyIle	140
OY	514	TTGGTGTCTTACAGCCAGGAGATCCGGGACATGTTCCACTGGCCCTCTTGTTGGAACACC	573
Db	141	LEULValPheTYrSerGIuGlnIleArGAspMetPheHisLeuArGleuPheValAspThr	160
OY	574	GACCTCCGACGTCAGGCTGTCTTGAAAGATCTCCGGACGTCGGCCGACGGGAGGAGCACTG	633
Db	161	AspSerAspAlaArgLeuSerArGArGValLeuArGAspValArGArGGLYArGAspLeu	180
OY	634	GACAGATTTCTGACGAGTACACCACTTCGTGAAGCCGGCCCTTGAGAGATTGTGCTG	693
Db	181	GIUGLIILLeuThrGlnTrThrThrPheValLYSProlaPheGluGlnPheCysLeu	200
OY	694	CCGACAAAGATATGCCATGTGATCATCCACGAGAGATGGACATATGCTTGGCATC	753
Db	201	ProThrLYSArTYrAlaAspValIleIleProArGGLYValaAspAsnMetValaIle	220
OY	754	AACCTGATCGTGACACATCCAGGACATTCGMAATGCTGACATCTGCAATGSCACACGA	813
Db	221	AsnLeuIleValGIuHisIleGlnAspIleLeuAsnGlyAspIleCysLYSTPrHisArg	240
OY	814	GAGAGGTCMAATGGGGGAGACTCAAGCGGACCTTTTCTGAGCGAGGAGACCACTTGGG	873
Db	241	GIYGLYSerAsnGlyArgSerTrLYrSArGTrhPheSerGIuProGlyAspHisProGly	260
OY	874	ATGCTGACCTCTGGCAAAAGGTCACATTTGGAGTCCAGACAGACACCCCAAC	924
Db	261	MetLeuThrSerGIyLYsArGSerHisLeuGlnSerSerSerArProHis	277
RESULT 5			
AAG64506 standard, Protein; 276 AA.			
XX	AA664506;		
XX	AC		
DT	02-OCT-2001 (first entry)		
XX			
DE	Human uridine kinase.		
XX			
KW	Human; uridine kinase; UK.		
XX			
OS	Homo sapiens.		
XX			
PN	CN1287172-A.		
XX			
PD	14-MAR-2001.		
XX			
PF	07-SEP-1999; 99CN-0118818.		
XX			
PR	07-SEP-1999; 99CN-0118818.		
XX			
PA	(UTFU-) UNIV FUDAN.		

PI Yu L, Zhao Y, Zhang H;
XX WPI; 2001-409529/44.
DR N-PSDB; AAH75355.
XX
PT Human uridine kinase and its coding sequence, preparation and application -
PS Claim 2; Page 15-16 (Disclosure); 20pp; Chinese.
CC The invention relates to human uridine kinase (UK).
XX
XX
SQ Sequence 276 AA;

Alignment Scores:
Pred. No.: 9,07e-104 Length: 276
Score: 1372.50 Matches: 266
Percent Similarity: 97.11% Conservative: 3
Best Local Similarity: 96.03% Mismatches: 7
Query Match: 46.26% Indels: 1
DB: 22 Gaps: 1

US-09-896-522-1 (1-1624) x AAG64506 (1-276)

QY 94 ATGGCTTCGGCGGAGCGAAGACTCGAGACCCCGCGGAGCCGACCTGCGCAC 153
Db 1 MetAlaSerAlaGlyGlyGlyGlyAlaArgAlaArgAla--GlyAlaAsnArgProHis 19

QY 154 CAGGCGCCCTTCGATAGGGGGTGAAGCGCGGCACTGCGAGCGGGAAGTGCAGCTGT 213
Db 20 GlnArgProPheLeuIleGlyValSerGlyGlyThrAlaSerGlyLysSerThrValCys 39

QY 214 GAGAGATCATGAGTTCCTGGGACAGAACGAGGTGGAACCGCGGACCGGAGGTGTC 273
Db 40 GluIleMetGluLeuLeuGlyGlnAsnGluValGluGlnArgHisGlyLysValVal 59

QY 274 ATCTGAGCCGAGGACAGGTTTACAAGGTCTTACCGGACGACGAGGAGCCAGCTTG 333
Db 60 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGluGlnLysAlaLysAlaLeu 79

QY 334 AAGAGACATGATTTTGAACCATCCAGATGCTTGTATATGATTTGATGACAGACT 393
Db 80 LysGlyIleIleThrAsnPheAspHisPheAspAlaPheAspHisPheAspHisPheAspThr 99

QY 394 CTGAGAACATCGTGAAGGGGCAAAACGATGAGGTGCGGACCTATGATTTTGTGACACAC 453
Db 100 LeuIleAsnIleValGluGlyLysThrValGluValProThrTyrAspPheValThrHis 119

QY 454 TCAAGTTTACCAAGACACGAGGTGCTTACCTCGGACGCTGTTCTTTGAGGCGATC 513
Db 120 SerArgLeuProGluThrThrValValTyrProAlaAspValValLeuPheGluGlyIle 139

QY 514 TTGGTGTCTTACAGCCAGAGATCCGGAGCATGTTTCAACCTGGGCGCTTCTGAGAACACC 573
Db 140 LeuValIleTyrSerGlnGluIleLeuArgAspPheHisLeuValGluLeuPheValAspThr 159

QY 574 GACTCCGAGCTCAGCTGTCTCGAAGATTCTCCGGGACGTCGCGCAGGAGGAGGACTG 633
Db 160 AspSerAspValArgLeuSerArgArgValLeuArgAspValArgArgGlyArgAspLeu 179

QY 634 GACGAGTTTCTGAGCGGATACACACCTTGTGAAAGCCGCTTCCAGAGATTCTGCTTG 693
Db 180 GluGlnIleLeuThrGlnTyrThrPheValLysProAlaPheGluGluPheCysLeu 199

QY 694 CCGACAAAGAAGTATGCGATGTGATCATCCACGAGAGGAGCAATATGTTGCATC 753
Db 200 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 219

QY 754 AACCTGATCGTGACGACATCCAGAGACATTTCTGATGATGACATCTGCAATGCGACGA 813
Db 220 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGluLysIleCysLysTyrPheHisArg 239

QY 814 GGAGGGTCCAAATGGGCGAGACTACAGCGGACCTTTTCTGAGCGAGGAGACCACTCTGG 873

Db 240 GlyGlySerAsnGlyArgThrTyrLysArgThrPheSerGluProGluAspHisPheGly 259
QY 874 ATGCTGACCTCTCGCAAAACGCTGACATTTTGAGTTCACGACGACAGACCCAC 924
Db 260 MetLeuThrSerGlyLysArgSerHisLeuGluSerSerSerArgProHis 276

RESULT 6
AA014412
ID AA014412 standard; Protein; 260 AA.
XX
XX AA014412;
XX
XX 02-MAY-2002 (first entry)
XX
XX
DE Protein of a human uridine kinase (UDK).
XX
XX
KW Human; uridine kinase; diagnostic assay; mutation detection; UDK;
KW probe; chromosome localization study; tissue expression; gene therapy;
KW antibody; vaccine; human ovarian cancer; immunological disorder;
KW human colon carcinoma; immunogen.
XX
XX Homo sapiens.
XX
XX WO200172963-A2.
XX
XX 04-OCT-2001.
XX
XX 27-MAR-2001; 2001WO-US09663.
XX
XX 27-MAR-2000; 2000US-0536647.
XX
XX (SMIK) SMITHKLINE BEECHAM CORP.
XX
XX Ho YS, Johnson RK;
XX
XX WPI; 2001-626259/72.
XX
XX N-PSDB; AAK98735.
XX
XX
PT Novel human uridine kinase polypeptides useful for treating cancers,
PT and to identify agonists and antagonists of the polypeptide useful for
PT treating conditions associated with uridine kinase imbalance -
XX
XX
PS Claim 3; Page 29-30; 31pp; English.
XX
XX The invention relates to newly identified human uridine kinase (UDK)
XX polypeptides and polynucleotides and methods for producing such
XX polypeptides by recombinant techniques. Also disclosed in the invention
XX are methods for utilizing uridine kinase polypeptides and polynucleotides
XX in diagnostic assays. The polynucleotides and polypeptides of the
XX invention may be used as diagnostic reagents by detecting mutations in an
XX associated gene. An array of oligonucleotide probes comprising the
XX uridine kinase polynucleotide sequence or fragments thereof can be
XX constructed to conduct efficient screening of genetic mutations, for
XX example. Detection of abnormally decreased or increased levels of
XX polypeptide or mRNA expression may also be used for diagnosing or
XX determining susceptibility of a subject to a disease of the invention.
XX The polynucleotide sequences of the invention can be used for chromosome
XX localization studies and tissue expression studies. The polypeptides of
XX the invention or fragments thereof may be used as immunogens to produce
XX antibodies. These antibodies may be employed to isolate or identify
XX clones expressing the polypeptide. The polypeptides and polynucleotides
XX of the invention can be used as a vaccine or in gene therapy to treat
XX diseases such as human ovarian cancer, human colon carcinoma, and
XX immunological disorders. This sequence represents the protein of a human
XX uridine kinase of the invention.
XX
SQ Sequence 260 AA;

Alignment Scores:
Pred. No.: 5,32e-103 Length: 260
Score: 1363.00 Matches: 260
Percent Similarity: 100.00% Conservative: 0

```
Best Local Similarity: 100.00%
Query Match: 45.94%
DB: 22
```

```

Mismatches: 0
Indels:     0
Gaps:       0

```

US-09-896-522-1 (1-1624) X AAO14412 (1-260)

OY		9	ATGGCTTCGGCGGAGGCACAAACTGTCCAGAACCCCGGACGGAGGCCAACCCTCCAC	153
Dd		1	MetalSerHisIedLylyLysrPySglusErrProAlarProGlnAlAspArgProHis	20
OY		154	CACCGGCCCTTCTGTATAGGGGTATGACGGCGGCACTGCCACGGGAACTGCACCTGTGT	213
Dd		21	GlnArgProPheLeuIleGlyValSerGIyGlyThrAlaSerGlyLysErrThrValCys	40
OY		214	GAGAAGATCATGGATTGCTGGGACACGAACAGGTGGAAACAGCCGGCACGGAACTGTC	273
Dd		41	GlnLysIleMetGluLeuLeuGluGlnAsnGlnValGluGlnArgGlnArgLysValVal	60
OY		274	ATCCTTAGCCAGACAGGTGTTCAAAGCTCTGACGGGACAGACAGAAAGCCAAAGCCCTTG	333
Dd		61	IleLeuSerGlnAspArgPheTyxLysValLysThrAlaGluGlnLysAlaLysValLeu	80
OY		334	AAAGCAGTACAATTGTAACCATCCAGATGCCCTTGTGAATAATGATTGATGACAGAGACT	393
Dd		81	LysGlyGlnItyxAsnPhenAsnIserProAsnAlaPheAsnAsnAspLeuMetHisArgThr	100
OY		394	CTGAAGACATCGTGGAGGGCAAAACGGTGGAGGTGCCGACCTATGATTGTTGTACACAC	453
Dd		101	LeuLysAsnIleValGlnUlyLysThrValGlnValProThrTyxAspPheValThnHis	120
OY		454	TCAAGGTACCAAGACSCACGGGGCTCAACCTCGGAGACGGGTCTGTGTAAGGGCATC	513
Dd		121	SerArgLeuProGluInthrThrValValTyxProAlaAspValValLeuPheGluGlyIle	140
OY		514	TTCGTGTTTACAGCCAGAGAGATCCGGGACATGTTCCACTGGCGCTTTTGTTGGAACCC	573
Dd		141	LeuValPheTyxSerGlnGlnIleArgAspMetPheHisIleuArgLeuPheValAspThr	160
OY		574	GATCCGAGCTCAGGCTGTCTCGAAGAGTTCTCCGGGACGTCCGCCGAGGGAGGGACCTG	633
Dd		161	AspSerAspValArgLeuSerArgValLeuArgAspValArgArgGLyArgAspLeu	180
OY		634	GACAGATTCTGACGAGTACACCAACTTCGGAAGCGGACCTTGAGAGATTCCTGCTG	693
Dd		181	GluGlnIleLeuThrGlnItyrThrThrPheValLysProAlaPheGluGlnPheCysLeu	200
OY		694	CCGACAAAGAGTATCCGATGTGATCATCCACAGAGAGTGGACATATATGTTGCCATC	753
Dd		201	ProThrLysLysTyxAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle	220
OY		754	AACCTGATCGTGGACACATCCAGGACATTTGAAATGTTGACATTTGTCAATGGCACCSGA	813
Dd		221	AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlnAspIleCysLysTyxPheAsnArg	240
OY		814	GGAGGGTCCAAATGGGAGGACTCAAGCGGACCTTTCTGAGCCAGGGACCAACCTGGG	873
Dd		241	GlyGlySerAsnMelTyxSerTyxLysArgInthrPheSerGluProGlyLysHisArgGly	260

AA014413	RESULT 7
AA014413	
ID	AA014413 standard; Protein; 277 AA.
XX	
AC	AA014413;
XX	
DT	02-MAY-2002 (first entry)
XX	
DE	Protein relating to a human uridine kinase (UDK) of the invention.
XX	
KW	Human; uridine kinase; diagnostic assay; mutation detection; UDK;
KW	probe; chromosome localisation study; tissue expression; gene therapy;
KW	antibody; vaccine; human ovarian cancer; immunological disorder;
XX	human colon carcinoma; immunogen.
XX	

OS Unidentified.
XX
PX WO200172963-A2.
XX
PD 04-OCT-2001.
XX
PF 27-MAR-2001; 2001WO-US09663.
XX
PR 27-MAR-2000; 2000US-0536647.
XX
PA (SMIK) SMITHKLINE BEECHAM CORP.
XX
PI Ho YS, Johnson RK;
XX
DR WPI; 2001-626259/72.
XX
XX
XX Novel human uridine kinase polypeptides useful for treating cancers, PT
PT to identify agonists and antagonists of the polypeptide useful for PT
PT treating conditions associated with uridine kinase imbalance -
XX
XX
PS Disclosure, Page 23, 31pp, English.

The invention relates to newly identified human uridine kinase (UDK) polypeptides and polynucleotides and methods for producing such polypeptides by recombinant techniques. Also disclosed in the invention are methods for utilizing uridine kinase polypeptides and polynucleotides in diagnostic assays. The polynucleotides and polypeptides of the invention may be used as diagnostic reagents by detecting mutations in an associated gene. An array of oligonucleotide probes comprising the uridine kinase polynucleotide sequence or fragments thereof can be constructed to conduct efficient screening of genetic mutations, for example. Detection of abnormally decreased or increased levels of polypeptide or mRNA expression may also be used for diagnosing or determining susceptibility of a subject to a disease of the invention. The polynucleotide sequences of the invention can be used for chromosome localization studies and tissue expression studies. The polypeptides of the invention or fragments thereof may be used as immunogens to produce antibodies. These antibodies may be employed to isolate or identify clones expressing the polypeptide. The polypeptides and polynucleotides of the invention can be used as a vaccine or in gene therapy to treat diseases such as human ovarian cancer, human colon carcinomas, and immunological disorders. This sequence represents the protein relating to a human uridine kinase (UDK) of the invention.

NOTE: The present sequence is stated as being the same as that shown as SEQ ID NO: 2 in the sequence listing of the specification. However, the sequences differ.

Sequence 277 AA;

Alignment Scores:	
Pred. No.:	7.03e-100
Score:	135.00
Percent Similarity:	95.67%
Best Local Similarity:	91.70%
Query Match:	44.66%
DB:	22
Length:	2477
Matches:	2540
Conservative:	11
Mismatches:	12
Indels:	0
Gaps:	0

US-09-896-522-1 (1-1624) x AAO14413 (1-277)

[illegible]

Db 61 IleuSerGlnAspCysPheTyrIysValLeuThrAlaGlnGlnLysAlaLysAlaLeu 80
QY 334 AAAGACAGTAACTTTTGGACATCCAGATGCTTTGTAATGATTGATCCAGAGACT 393
Db 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisLysThr 100
QY 394 CTGAAGAATCATCTGGAGGCAAAAAGGTGGAGGTCCGACTATGATTGTGACACAC 453
Db 101 LeuLysAsnIleValGlnGlyLysThrValGlnValProThrTyrAspPheValThrHis 120
QY 454 TCAAGTTACAGAGACACCGAGTGTACCTCGGAGCGCTGTCTGTTGAGGAGCATC 513
Db 121 SerArgLeuProGlnThrThrValValTyrProAlaAspValValLeuPheGlnGlyIle 140
QY 514 TTGGTGTTCACAGCCAGAGATCCGGACATGTTCCACTGCGCTCTTGTGAGACAC 573
Db 141 LeuValPheTyrThrGlnGlnIleAlaArgAspMetPheHisLeuAlaArgLeuPheValAspThr 160
QY 574 GACTCCGACGTCAGGCTGTCTCGAAGATTTCTCCGGACGTGCCCGGAGGAGGACCTG 633
Db 161 AspSerAspValAlaArgLeuSerArgArgValLeuAlaArgAspValGlnArgGlyArgAspLeu 180
QY 634 GAGAGATTCTGACGACATACACACCTTGTAAGCGGCTTCCAGAGATTGCTGCTG 693
Db 181 GlnGlnIleLeuThrGlnTyrThrAlaPheValLysProAlaPheGlnGlnPheCysLeu 200
QY 694 CCGACAAAGAATATGCCGATGTATCATCCACAGAGATGACAAATATGTTGCCATC 753
Db 201 ProThrLysLysTyrAlaAspValIleIleProArgGlyValAspAsnMetValAlaIle 220
QY 754 AACCTGATCTGCGACATCCAGACATTTGTAATGTCATCTGCAATGGCACCGA 813
Db 221 AsnLeuIleValGlnHisIleGlnAspIleLeuAsnGlnAspLeuCysLysArgHisArg 240
QY 814 GGAAGGTCCATGGGCGGAGCTACAGAGCACTTTCTGAGCGGAGGACACCTGGG 873
Db 241 GlyGlyProAsnGlnArgAsnHisLysArgThrPheProGlnProGlnAspHisProGly 260
QY 874 ATGCTGACCTCTGGCAAAAGGTCAATTTGAGAGTCCAGACAGACCCAC 924
Db 261 ValLeuAlaThrGlyLysArgSerHisLeuGlnLysSerSerArgProHis 277
RESULT 8
ID ABB89353 standard; Protein; 190 AA.
AC ABB89353;
DT 24-MAY-2002 (first entry)
DE Human polypeptide SEQ ID NO 1729.
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KM antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
KM vulnerable; anticorrosive; antibacterial; antifungal; antiparasitic;
KM cardiac; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein.
OS Homo sapiens.
PN MO200190304-A.
PD 29-NOV-2001.
PF 18-MAY-2001; 2001WO-US16450.
PR 19-MAY-2000; 2000US-205515P.
PA (HUMA-) HUMAN GENOME SCI INC.
PI Birse CE, Rosen CA;
XX WPI; 2002-122018/16.

DR N-PSDB; ABL89762.
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
PT prevention of neural, immune system, muscular, reproductive,
PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
PT disorders -
XX
PS Claim 11; SEQ ID NO 1729; 2081pp + Sequence Listing; English.
XX
CC The invention relates to novel genes (ABL9449-ABL90853) and proteins
CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
CC medical conditions e.g. by protein or gene therapy. The genes are
CC isolated from a range of human tissues disclosed in the specification.
CC The nucleic acids, proteins, antibodies and (ant)agonists are useful
CC in the diagnosis, treatment and prevention of: (a) cancer, e.g. breast
CC and ovarian cancer and other cancers of the adrenal gland, bone, bone
CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune
CC hemolytic anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's
CC disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases such as viral, bacterial, fungal
CC and parasitic infections.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at fcp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 190 AA;
Alignment Scores:
Pred. No.: 2,56e-69 Length: 190
Score: 951.00 Matches: 187
Percent Similarity: 97.91% Conservative: 0
Best Local Similarity: 97.91% Mismatches: 3
Query Match: 32.05% Indels: 1
DB: Gaps: 0
US-09-896-522-1 (1-1624) x ABB89353 (1-190)
QY 94 ATGGCTTCGGCGGAGGCGAAGCTGCGAGGCCCGCGGAGCGGACCGTCCGCAC 153
Db 1 MetAlaSerAlaGlyLysAspCysGlnSerProAlaProGlnAlaAspArgProHis 20
QY 154 CAGCGGCCCTTCTGTAAGGGGTGAGCGGCGGCACTCCAGCGGAGTGCACCTGTGT 213
Db 21 GlnArgProPheLeuIleGlyAlaSerGlyThr***SerGlyLysSerThrValCys 40
QY 214 GAGAAGATCATGGAGTTGCTGGGACAGAAACGAGTGGAAACGCGGACGGAAAGTGTG 273
Db 41 GlnLysIleMetGlnLeuLeuGlnGlnAsnGlnValGlnGlnArgGlnArgLysValVal 60
QY 274 ATCTGAGCGAGACAGGTTCATACAGGTCTCGGACGCGAGAGAAAGCCAGGCTTG 333
Db 61 IleLeuSerGlnAspArgPheTyrLysValLeuThrAlaGlnGlnLysAlaLysAlaLeu 80
QY 334 AAAGACAGTAACTTTTGGACATCCAGATGCTTTGTAATGATTGATCCAGAGACT 393
Db 81 LysGlyGlnTyrAsnPheAspHisProAspAlaPheAspAsnAspLeuMetHisLysThr 100
QY 394 CTGAAGAATCATCTGGAGGCAAAAAGGTGGAGGTCCGACTATGATTGTGACACAC 453
Db 101 LeuLysAsnIleValGlnGlyLysThrValGlnValProThrTyrAspPheValThrHis 120
QY 454 TCAAGTTACAGAGACACCGAGTGTACCTCGGAGCGCTGTCTGTTGAGGAGCATC 513
Db 121 SerArgLeuProGlnThrThrValValTyrProAlaAspValValLeuPheGlnGlyIle 140
QY 514 TTGGTGTTCACAGCCAGAGATCCGGACATGTTCCACTGCGCTCTTGTGAGACAC 573
Db 141 LeuValPheTyrThrGlnGlnIleAlaArgAspMetPheHisLeuAlaArgLeuPheValAspThr 160
QY 574 GACTCCGACGTCAGGCTGTCTCGAAGATTTCTCCGGACGTGCCCGGAGGAGGACCTG 633

```

Db      161 AspSerAspValArgLeuSerArgArg***LeuArgAspValArgArgGlyArgAspLeu 180
QY      634 GAGCAGATTCTGACGCGAGTACACCACTTCG 664
        |||||
Db      181 GluGlnIleuThrHis-ThrProProSer 190

RESULT 9
AAM38694
ID      AAM38694 standard; Protein; 261 AA.
XX
AC      AAM38694;
XX
DT      22-OCT-2001 (first entry)
XX
DE      Human polypeptide SEQ ID NO 1839.
XX
KW      Human; nocotropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW      peripheral nervous system; neuropathy; central nervous system; CNS;
KW      Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW      amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW      chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW      leukaemia.
XX
OS      Homo sapiens.
XX
PN      MO200153312-A1.
XX
PD      26-JUL-2001.
XX
PF      26-DEC-2000; 2000MO-US34263.
XX
PR      21-JAN-2000; 2000US-0488725.
PR      25-APR-2000; 2000US-0552317.
PR      09-JUL-2000; 2000US-0598042.
PR      19-JUL-2000; 2000US-0620312.
PR      03-AUG-2000; 2000US-0653450.
PR      14-SEP-2000; 2000US-0662191.
PR      19-OCT-2000; 2000US-0693036.
PR      29-NOV-2000; 2000US-0727344.
XX
PA      (HYSE-) HYSEQ INC.
XX
PI      Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI      Wang J, Wang Z, Wehman T, Xu C, Xue AJ, Yang Y, Zhang J;
PI      Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX
DR      WPI; 2001-442253/47.
DR      N-PSDB; AAI57850.
XX
PT      Novel nucleic acids and polypeptides, useful for treating disorders
PT      such as central nervous system injuries -
PS      Example 3; SEQ ID NO 1839; 10078bp; English.
XX
CC      The invention relates to human nucleic acids (AA157798-AA161369) and
CC      the encoded polypeptides (AAM38642-AAM42213) with nocotropic,
CC      immunosuppressant and cytostatic activity. The polynucleotides are useful
CC      in gene therapy. A composition containing a polypeptide or polynucleotide
CC      of the invention may be used to treat diseases of the peripheral nervous
CC      system, such as peripheral nervous injuries, peripheral neuropathy and
CC      centralised neuropathies and central nervous system diseases, such as
CC      Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
CC      lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
CC      utilisation of the activities such as: Immune system suppression,
CC      Actin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
CC      and thrombolytic activity, cancer diagnosis and therapy, drug screening,
CC      assays for receptor activity, arthritis and inflammation, leukaemias and
CC      C.N.S disorders.
CC      Note: The sequence data for this patent did not form part of the printed
CC      specification.
XX
SQ      Sequence      261 AA;

```

```

Alignment Scores:
Pred. No.:      7.05e-68      Length:      261
Score:          934.00      Matches:      189
Percent Similarity: 79.41%      Conservative: 27
Best Local Similarity: 69.49%      Mismatches: 40
Query Match:    31.48%      Indels:      16
DB:            22      Gaps:      4

US-09-896-522-1 (1-1624) x AAM38694 (1-261)
QY      115 GACTGCGAGAGCCCGCGCGGAGGCGGACGTCGCGAC---CAGCGGCCCTTCGATA 171
        |||
Db      4 AspSerGluGlnThrLeuGlnIleuHisGlnGlnProAsnGlyGlyGluProHeuLeuIle 23
QY      172 GGGGTGAGGCGGCGGACCTGCCAGCGGAAAGTCGACCGTGTGTGAGAAAGATCATGAGTTG 231
        |||||
Db      24 GlyValSerGlyGlyThrAlaSerGlyLysSerSerValCysAlaIleValGlnLeu 43
QY      232 CTGGGACAAAGAGGTGGAACAGCGGCGCGGAAAGTGTGATCTGATCGCCAGGACAG 291
        |||||
Db      44 LeuGlyGlnAsnGluValAlaAspTyrArgGlnLysGlnValAlaIleLeuSerGlnAspSer 63
QY      292 TTCTACAGGTCCTGACGCGGACGAGCAAGGCGCAAGGCGCTTGAAGAAGACATATT 351
        |||||
Db      64 PheTyrArgValLeuThrSerGluGlnLysAlaLysAlaLeuLysGlyLysPheAsnPhe 83
QY      352 GACCATCGAGATGCTCTTGATATGATTGATGACACGAGACTCTGAAAGACATCTGGAG 411
        |||||
Db      84 AsnHisProAspAlaPheAspAsnGlnLeuIleLeuLysThrLeuLysGlnIleThrGln 103
QY      412 GCGAAAGCGGTGAGGTGGCGGACCTGATGATTTGTGACACATCAAGGTTACAGAGACC 471
        |||||
Db      104 GlyLysThrValGlnIleProValTyrAspPheValSerHisSerTyrGlnGluGlnTyr 123
QY      472 ACGGTGTCTTACCGTCGCGGACGAGTGTGTTGTGAGGGCATTTGGTGTCTTACAGCCAG 531
        |||||
Db      124 ValThrValTyrProAlaAspValValLeuPheGluGlyIleLeuAlaPheTyrSerGln 143
QY      532 GAGATCGGAGACATGTTCCACCTGCGGCTTCTGTGTGACACCGACTCCGACGTCAGGCTG 591
        |||||
Db      144 GluValArgAspLeuPheGlnIleMetLysLeuPheValAspThrAspAlaAspThrArgLeu 163
QY      592 TTCTGAAAGGTTCTCCGGGACGTG---CGCCGAGGAGGAGGACCTTGACGACATTTGCAAG 648
        |||||
Db      164 SerArgArgValLeuArgAspIleSerGlnArgGlyArgAspLeuGlnIleLeuSer 183
QY      649 CAGTACACCACTTCGCGTGAAGCGGCGCTTCAGAGAGTTGCTGCGCGCAAGAAAGTAT 708
        |||||
Db      184 GlnTyrIleThrPheValLysProAlaPheGlnGluPheCysLeuProThrLysLysTyr 203
QY      709 GCCGATGTGATCATCCACGAGAGGTGACAATATGATGTTGCCATCAACTGATCTGACAG 768
        |||||
Db      204 AlaAspValIleIleProArgGlyAlaAspAsnLeuValAlaIleAsnLeuIleValGln 223
QY      769 CACATCCAGAGATCTCTGAATGTGACATCTGCCAATGSCACCCGAGAGGTCCAATGGG 828
        |||||
Db      224 HisIleGlnAspIleLeuAsnGly-----Gly 232
QY      829 CGAGCTTACACCGGACCTTTCTGAGCCAGGGGACCACTTGAGATGCTGACCTGTGGC 888
        |||||
Db      233 ProSerLysArgGlnThr-----AsnGlyCysLeuAsnGlyTyrThrProSerArg 249
QY      889 AAACGGTCACATTGGAGTCCAGACAGACAGACCCAC 924
        |||||
Db      250 LysArgGlnAlaSerGlnLysSerSerArgProHis 261
        |||||

RESULT 10
AAB73494
ID      AAB73494 standard; Protein; 261 AA.
XX
AC      AAB73494;
XX

```

DT 31-JUL-2001 (first entry)
 XX
 DE Human transferrase HTFS-1, SEQ ID NO:1.

XX Human transferrase; HTFS, agonist; antagonist; cellular signalling;
 XX proliferation; cell proliferative disorder; immune disorder;
 XX atherosclerosis; hepatitis; psoriasis; cancer; tumour;
 XX inflammation; AIDS; Addison's disease; allergy; asthma; anaemia;
 XX cirrhosis; Crohn's disease; atopic dermatitis; diabetes mellitus;
 XX multiple sclerosis; rheumatoid arthritis; pancreatitis;
 XX systemic lupus erythematosus; thrombocytopenia; ulcerative colitis;
 XX haemodialysis; extracorporeal circulation; trauma; transgenic animal;
 XX gene therapy; drug screening.

OS Homo sapiens.
 XX
 XX WO200132888-A2.

XX 10-MAY-2001.

XX 02-NOV-2000; 2000WO-US30485.

XX 04-NOV-1999; 99US-0163595.

XX (INCY-) INCYTE GENOMICS INC.

XX Tang YT, Yue H, Hillman JL, Lal P, Bandman O, Patterson C;
 PI Shih IL, Azimzai Y, Lu DAM, Baughn MR;

XX WPI; 2001-328796/34.

XX N-PSDB; AAH23801.

PT Human transferrase polypeptides and polynucleotides useful for
 PT diagnosis, prevention and treatment of cell proliferative and immune
 PT system disorders and for identifying agonists and antagonists -

PS Claim 1; Page 103-104; 157pp; English.

XX Sequences AAB73494-AAB73535 represent novel human transferrase proteins
 HTFS-1 to HTFS-42, and sequences AAH23801-AAH23842 represent cDNAs
 encoding them. The proteins play important roles in the regulation of
 cellular signalling and proliferation. The HTFS proteins are useful for
 screening compounds for their effectiveness as agonists or antagonists of
 transferrase activity, or for compounds that specifically bind to an HTFS
 protein or which modulates the activity of an HTFS protein.
 CC Pharmaceutical compositions comprising an HTFS protein, HTFS
 CC agonist or antagonist, or genetic construct encoding an HTFS
 CC protein are useful for treating a disease or condition associated
 CC with decreased or increased expression of functional HTFS. Disorders
 CC which may be treated using such compositions include cell proliferative
 CC disorders and immune disorders. For example, diseases which may be
 CC treated include atherosclerosis, hepatitis, psoriasis, cancer (including
 CC breast, bladder, bone marrow, brain and uterus cancer), inflammation,
 CC AIDS, Addison's disease, allergies, asthma, anaemia, cirrhosis, Crohn's
 CC disease, atopic dermatitis, diabetes mellitus, multiple sclerosis,
 CC rheumatoid arthritis, pancreatitis, systemic lupus erythematosus,
 CC thrombocytopenia, and ulcerative colitis. They may also be used to treat
 CC complications of cancer, haemodialysis, extracorporeal circulation,
 CC trauma and haemopoietic cancer, including lymphoma, leukemia and
 CC myeloma. Polynucleotides encoding HTFS proteins are useful for creating
 CC transgenic animals to model human diseases, for diagnostic purposes and
 CC to generate hybridisation probes useful in mapping the naturally
 CC occurring genomic sequences. HTFS, and its catalytic or immunogenic
 CC fragments are useful for screening libraries of compounds in a variety of
 CC drug screening techniques. Antibodies which specifically bind HTFS may be
 CC used for the diagnosis of disorders associated with the expression of
 CC HTFS, or in assays to monitor patients being treated with HTFS or
 CC agonists, antagonists or inhibitors of HTFS. The present sequence
 XX represents an HTFS protein of the invention.

XX Sequence 261 AA;

Alignment Scores:

Pred. No.: 7,05e-68 Length: 261
 Score: 934.00 Matches: 189
 Percent Similarity: 79.41% Conservative: 27
 Best Local Similarity: 69.49% Mismatches: 40
 Query Match: 31.48% Indels: 16
 DB: 22 Gaps: 4

US-09-896-522-1 (1-1624) X AAB73494 (1-261)

QY	115	GACTGCGAGAGCCCGCGGAGGCCGACCCGTCGAC--CAGCGGCGCTTCCTGATA	171
DB	4	AspSerGluInThrLeuGlnAsnHISgInGlnProAsnGlyGlyGluProPheLeuIle	23
QY	172	GGGATGACGGCGGACCTGCGACGCGGAGTGCACCGTGTGTGAGAAAGCATCATGAGTT	231
DB	24	GlyValSerGlyGlyThrAlaSerGlyLysSerValCysAlaIleValIleLeuSerGlnAspSer	43
QY	232	CTGGGACAGAACGAGGTGAAACAGCGGACGCGGAGGTGTCATCTTGACCGACGACG	291
DB	44	LeuGlyGlnAsnGluValAspTyrArgGlnLysGlnValValIleLeuSerGlnAspSer	63
QY	292	TTCCTACAGGTCTCTGACGCGACGACGAGGAGCCGAGGCTTGAAAGACATGATTTT	351
DB	64	PheTyrArgValLeuThrSerGluGlnLysAlaLysAlaLeuLysGlyGlnPheAsnPhe	83
QY	352	GACCATCAGATGCTTGTGATTAATGATTGATGACACGACCTGGAAGAACGTCGAG	411
DB	84	AspHisProAspAlaPheAspAsnGluLeuIleLeuLysThrLeuLysGluIleTrpGlu	103
QY	412	GGCAAAACGGTGGAGGTGCGGACCTATGATTTGTGACACACTCAAGATTACCGAGACC	471
DB	104	GlyLysThrValGlnIleProValTyrAspPheValSerHisSerArgLysGluIleTrp	123
QY	472	ACGGTGTCTACCTTCGCGACCGTCGTGTTGTTGAGGCGATCTTGTTCTACACCCAG	531
DB	124	ValThrValTyrProAlaAspValValLeuPheGlnGlyIleLeuAlaPheTyrSerGln	143
QY	532	GAGATCCGGGACATGTTCCACCTGCGCTCTTGTGTGACACGACCTCGACGTCG	591
DB	144	GluValArgAspLeuPheGlnMetLysLeuPheValAspThrAspAlaAspThrArgLeu	163
QY	592	TCTGGAAGATTCTCGGGAAGTGTG---CGCGAGGAGGAGGACCTTGAGGATTCGTGACG	648
DB	164	SerArgArgValLeuArgAspIleSerGluArgGlyArgAspLeuGlnIleLeuSer	183
QY	649	CAGTACACACCTTCTGTGAAGCCGCGCTTCGAGGATTCGCTCGCCGACAAAGAGTAT	708
DB	184	GlnTyrIleThrPheValLysProAlaPheGlnGluPheCysLeuProThrLysLysTyr	203
QY	709	GCGGATGTGATCTCCACGAGGAGTGACCAATATGTTGCCATCAACCTGATTCGCGACG	768
DB	204	AlaAspValIleIleProArgGlyAlaAspAsnLeuValAlaIleAsnLeuIleValGln	223
QY	769	CACATCCAGACATCTTGTAATGTGTGACATCTCGAAATGCGACCGGAGGTCATGCGG	828
DB	224	HisIleGlnAspIleLeuAsnGly-----Gly	232
QY	829	CGGAGCTACAGCGGACCTTTCTGACGAGGAGACCCCTGGATGCTGACCTTGCG	888
DB	233	ProSerLysArgGlnThr-----AsnGlyCysLeuAsnGlyLysThrProSerArg	249
QY	889	AAACGTCATATTGGATTCGACGACGACGACCCAC	924
DB	250	LysArgGlnAlaSerGluSerSerSerArgProHis	261
RESULT 11			
ID	ABP41393	standard; Protein; 337 AA.	
AC	ABP41393;		
XX			
DT	22-AUG-2002	(first entry)	
XX			

DE Human ovarian antigen HOELP29, SEQ ID NO:2525.
 XX
 KM Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
 KM ovarian cancer; breast cancer; tumour; reproductive system disorder;
 KM infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
 KM PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
 KM inflammatory condition; immune disorder; blood disorder;
 KM cardiovascular disorder; respiratory disorder; neurological disorder;
 KM gastrointestinal disorder; urinary system disorder; drug screening;
 KM gene therapy; chromosome mapping; forensic analysis;
 KM antibody preparation; cytosolic; immunomodulatory; neuroprotective;
 KM antiinflammatory; gynaecological; reproductive; chromosome 1p32.
 OS Homo sapiens.
 XX
 XX MO200200677-A1.
 XX
 PD 03-JAN-2002.
 XX
 PF 07-JUN-2001; 2001MO-US18569.
 XX
 PR 07-JUN-2000; 2000US-209467P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Birse CE, Rosen CA;
 XX
 DR WPI, 2002-147878/19.
 XX
 DR N-PDB; ABO54470.
 XX
 PT Isolated nucleic acid molecules encoding novel ovarian polypeptides,
 PT useful in the prevention, treatment and diagnosis of cancer (e.g.
 PT ovarian cancer), immune disorders, cardiovascular disorders and
 PT neurological diseases -
 XX
 PS Claim 11: SEQ ID NO 2525; 2922bp; English.
 XX
 XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
 CC ABP43228) and to cDNAs encoding them (ABO54131-ABO56305), and also
 CC encompasses polypeptides 90% identical and polynucleotides 95% identical
 CC to the sequences of the invention. The invention additionally relates to
 CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosting or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents a human ovarian antigen of the invention.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 XX
 XX
 SQ Sequence 337 AA;

Alignment Scores:
 Pred. No.: 1.64e-67 Length: 337
 Score: 930.00 Matches: 201
 Percent Similarity: 73.86% Conservative: 25

Best Local Similarity: 65.69% Mismatches: 57
 Query Match: 31.34% Indels: 23
 DB: 23 Gaps: 6
 US-09-896-522-1 (1-1624) x ABP41393 (1-337)
 QY 12 TCAGACCTGGGCGCTGGGCGGCGCGCGCGGAGGAGGCGCGGAGCCG 71
 Db 54 SerhpaAlGIYAlaGly--GlyGluArgAlaSerValArgThrGlySerGly----- 70
 QY 72 ATGCGCGGAGCGGAGCGCGGAGATGCTTCGCGGAGCGGAGGAGAGCTGCGAGACCCCGC 131
 Db 71 ---ArgArgGlyGlyAlaSerhisiGlyArgGlyGluArgAlaAspProAlaGluProPro 89
 QY 132 GCCGAGGCGCGACCGTCGACACGAGCGCGCTTCCTGATAGGGGTGAGCGCGCGCTGC 191
 Db 90 AlaAlaGlnArg-----ArgArgAlaLeuProTyArgArgHis-GlyGlyThrAl 106
 QY 192 CAGCGGAGATCGACCGTGTGAGAGATCATGAGAGTGTGCGACAGAGAGGTGAG 251
 Db 106 AserGlyYpSerSerValCysAlaAlaYsileValGlnLeuLeuGlyGlnhnglValhAs 126
 QY 252 ACAGCGGACGCGAGAGGTGATCCTCGACGAGACAGATTCTACAGGTCTCGACGCGC 311
 Db 126 pTyArgGlnGlyGlnValValIleLeuSerGlnAspSerPheTyArgValLeuThrse 146
 QY 312 AGAGCGAGAGCGCGAGGCTTGAGAGAGATTCATTTGACATCCAGATGCTTTGA 371
 Db 146 rGluGlnYsAlaYsAlaLeuLys**GlnPhehAsnPheAspHisProAspAlaPheAs 166
 QY 372 TATATGTTGATGACGACGAGACTCTGAGAAACATCCGAGGCGCAAAACGATGAGGTGCC 431
 Db 166 pAsnGlnLeuIleLeuYsThrLeuYsGluIleThrGlnGlyYsThrValGlnIleThr 186
 QY 432 GACCTATGATTTTGTGACACACTCAAGTTTACAGAGACCAAGTGTCTACCTCGAG 491
 Db 186 oValTyAspPheValSerHisSerArgYsGlnGluThrValThrValTyProAlaAs 206
 QY 492 CGTGTCTCTGTTGAGGCGATCTGTGTTCTACAGCCGAGAGATCCGGAGATGTTCCA 551
 Db 206 pAlaValLeuPheGlnGlyYsIleLeuAlaPheTySerGlnGlnValhArgAspLeuPheG 226
 QY 552 CCGGCGCTCTCGTGGACACCGACTCGGACGCGGCTGTCTCGAAGGTTCTCGGGA 611
 Db 226 hMetYsLeuPheValAspThrAspAlaAspThrArgLeuSerArgValhLeuArgAs 246
 QY 612 CGTG--CCCGAGGAGGAGACTGTGAGAGATTCAGCAGTACACCACTTCGTGAA 668
 Db 246 pIleSerGluArgGlyArgAspLeuGlnIleLeuSerGlnTyIleThrPheValY 266
 QY 669 GCCGCGCTTGGAGAGTTCGCTGCGCAAAAGATATGCCGATGTATATCCACG 728
 Db 266 sProAlaPheGlnGluPheCysLeuPheProThrYsTyAlaAspValIleIleProAr 286
 QY 729 AGGAGTGCACATPAGTGTGACATCAACTGATCGGACGACATTCAGACATTTGAA 788
 Db 286 gGlyAlaAspAsnLeuValAlaIleSerLeuIleValGlnhSileGlnAspIleLeuAs 306
 QY 789 TGGTGCATCTGCMAATGACACCGAGAGGTTCATATGCGGCGGAGTCAAGCGGACCTT 848
 Db 306 ngly-----GlyProSerLeuArgGlnThr-- 314
 QY 849 TTCTGACGAGGAGACACCTCGGATGCTGACCTTGCAAAACGTCACATTTGGAGTC 908
 Db 315 -----AsnglyCySLeuAsnglyTyThrProSerArgYsArgGlnAlaSerGlySe 332
 QY 909 CAGCAGACAGACCCAC 924
 Db 332 rSerSerArgProHis 337
 RESULT 12
 AAB56582
 ID AAB56582 standard; Protein; 337 AA.

XX	AAB56582;
AC	
XX	
DT	13-MAR-2001 (first entry)
XX	
DE	Human prostate cancer antigen protein sequence SEQ ID NO:1160.
XX	
KW	Humans; prostate cancer; prostate cancer antigen; detection; diagnosis; neuroprotective; cytoskeletal; cardioprotective; immunomodulatory; muscular; vulnerable; gastrointestinal; nephrotoxic; antiinfective; gynaecological, KW antibacterial; gene therapy; neural; immune; reproductive; renal; KW gastrointestinal; pulmonary; cardiovascular; proliferative disorder; wound; infectious disease.
XX	
OS	Homo sapiens.
XX	
PN	WO20005174-A1.
XX	
PD	21-SEP-2000.
XX	
PF	08-MAR-2000; 2000MO-US05988.
XX	
PR	12-MAR-1999; 99US-0124270.
XX	
PA	(HUMA-) HUMAN GENOME SCI INC. (ROSE/) ROSEN C A.
XX	
PI	Rosen CA, Ruben SM;
XX	
DR	WPI, 2000-587513/55.
N-PSDB:	AAFI5785.
XX	
PT	Prostate cancer associated gene sequences, referred to as prostate cancer antigens, useful for treatment, prevention, and diagnosis of disorders such as prostate cancer -
XX	
PS	Claim 11; Page 1566-1567; 2338bp; English.
XX	
CC	AAFI5566 to AAFI6505 encode the human prostate cancer associated proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC	The prostate cancer antigens can have neuroprotective, cytostatic, cardioactive, immunomodulatory, muscular, vulnary, gastrointestinal, nephrotoxic, antiinfective, gynaecological and antibacterial activities, and can be used in gene therapy. The prostate cancer antigen polynucleotides may be used for detection of prostate cancer, chromosome identification, as chromosome markers, and for numerous other diagnostic or research purposes. The prostate cancer antigens may be used to treat disorders such as neural, immune, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, and proliferative disorders, wounds, and infectious diseases. AAFI6506 to AAFI6514 to AAB57303 represent sequences used in the exemplification of the present invention.
CC	
XX	
SQ	Sequence 337 AA;
Alignment Scores:	
Pred. No.:	4,21e-67
Score:	925.00
Percent Similarity:	73.53%
Best Local Similarity:	65.36%
Query Match:	31.18%
DB:	21
	Gaps: 6

[illegible]

QY	132	GGCCGAGGCGCAACCGTCCGACACCGAGCGCCCTTCTCGATAGGGGTGACGGCGGACCTGC	191
Db	90	AlaAlaGlnArg-----ArgArgAlaLeuProTyrArgArgHis-GlyGlyThrAla	106
QY	192	CAGCGGGAAGTGCACCGTGTGTGAGAAAGATCATGAGTGTCTGGGACAGAAACAAGGTGGA	251
Db	106	AserIlyysSerSerValCysAlaAlaValAlaGlnLeuLeuGlyGlnAsnGluValAs	126
QY	232	ACAGCGGACAGCGGAAGTGGTGCATCCTGAGCCAGGACAGATTCTACAGAGTCTCTGACGG	311
Db	126	PtyArgGlnIlysGlnValValIleuSerGlnAspSerPheTyrArgValLeuThrSe	146
QY	312	AGAGGAGAAGGCGCAAGCCCTTGAAGAAGACAGTACATTTTGACACATCCGATGCCCTTGA	371
Db	146	rgIuGlnIlysaIyAlaAlaLeuLys**GlnPheAsnPheAspHisProAspAlaPheAs	166
QY	372	TAATGATTGTGATGCACAGACTCTGAAGAACAATCGGAGGGGCAAAACGGTGAAGGTGC	431
Db	166	pAsnGlu***IleuLysThrLeuLysGluIleThrGlnGlyLysThrValGlnIleer	186
QY	432	GACCTATGATTTGTGACACACTCAAGGTTTACCAAGACACACGGTGTCTACCCCTGCGA	491
Db	186	oValItyrAspPheValSerHisSerArgLysGlnGluThrValThrValItyrProAlaAs	206
QY	492	CGTGTGTCTGTTTGAAGGCATCTTGGTGTCTTACACGCCAGGAGATCCGGACATGTTCCA	551
Db	206	pValValIleuPheGlnGlyIleLeuAlaPheTyrSerGlnGluValArgAspLeuPheG1	226
QY	552	CCTGCGCCTCTTCTGAGACACCGACTCCGACGCTGCTCGAAGAGTCTCCGGGA	611
Db	226	nMetLysLeuPheValAspThrAspAlaAspThrArgLeuSerArgValIleuArgAs	246
QY	612	CGTG---CGCGAGGAGAGGACCTTGAGAGATTTGTACGACAGTACACCACTTCGTGAA	668
Db	246	pIleSerGlnArgGlyArgAspLeuGlnGlnIleuSerGlnTyrIleThrPheValIy	266
QY	669	GCCGCGCTTTCGAGAGATTCTGCTGCCGACAAAGAATATGCCATGTATCATCCCAAG	728
Db	266	sProIaPheGlnGluPheCysLeuProThrLysLysTyrAlaAspValIleIleProAr	286
QY	729	AGGAGTGACAATATGTTGGCATCAACTGATCGTGAGCAGCATCCAGACATTTGTGAA	788
Db	286	gGlyAlaAspAsnLeuValAlaIleAsnLeuIleValGlnHisIleGlnAspIleuAs	306
QY	789	TGGTGACATCTGCAAATAGGCAACGAGAGGGTCCATGGGCGGACCTACAGCGAGCTT	848
Db	306	nGly-----GlyProSerLysArgGlnThr--	314
QY	849	TTCTGAGCGAGGAGCACACCTGGGATGCTGACTCTGGCAAAACGGTACATTTGGAGTC	908
Db	315	-----AsnGlyCysLeuAsnGlyTyrThrProSerArgLysArgGlnAlaSerGluSe	332
QY	909	CAGCAGCAGACCCAC	924
Db	332	rSerSerArgProHis	337
RESULT 13			
ID	AA440480	standard; Protein; 335 AA.	
AC	AA440480;		
XX			
DT	22-OCT-2001	(first entry)	
DE		Human polypeptide SEQ ID NO 5411.	
XX			
KW		Human; noctropic; immunosuppressant; cyostatic; gene therapy; cancer;	
KW		peripheral nervous system; neuropathy; central nervous system; CNS;	
KW		Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;	
KW		amyotrophic lateral sclerosis; Shy-Drager syndrome; chemoclastic;	
KW		chemokineic; thrombolytic; drug screening; arthritis; inflammation;	
XX		leukaemia.	
XX			

XX Human, nootropic, immunosuppressant, cytostatic; gene therapy; cancer;
 KW peripheral nervous system; neuropathy; central nervous system; CNS;
 KW Alzheimer's, Parkinson's disease, Huntington's disease; haemostatic;
 KW amyotrophic lateral sclerosis; Shy-Drager syndrome; chemotactic;
 KW chemokine; thrombolytic; drug screening; arthritis; inflammation;
 KW leukaemia.
 XX

XX	Human sapiens.
XX	W0200153312-AL.
XX	26-JUL-2001.
XX	26-DEC-2000; 2000MO-US34263.
XX	21-JAN-2000; 2000US-0488725.
XX	25-APR-2000; 2000US-0552317.
XX	09-JUL-2000; 2000US-0598042.
XX	19-JUL-2000; 2000US-0620312.
XX	03-AUG-2000; 2000US-0653450.
XX	14-SEP-2000; 2000US-0662191.
XX	19-OCT-2000; 2000US-0693036.
XX	29-NOV-2000; 2000US-0727344.
XX	(HYSEQ-) HYSEQ INC.
XX	Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D,
XX	Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J,
XX	Zhao QA, Zhou P, Goodrich R, Drmanac RT;
XX	WPI: 2001-442253/47.
XX	N-PSDB; AA159636.
XX	Novel nucleic acids and polypeptides, useful for treating disorders
XX	such as central nervous system injuries -
XX	Example 2; SEQ ID NO 5411; 10078bp; English.
XX	The invention relates to human nucleic acids (AA157798-AA161369) and
XX	the encoded polypeptides (AA138642-AA142213) with neurotropic,
XX	immunosuppressant and cytoskeletal activity. The polynucleotides are useful
XX	in gene therapy. A composition containing a polypeptide or polynucleotide
XX	of the invention may be used to treat diseases of the peripheral nervous
XX	system, such as peripheral nervous injuries, peripheral neuropathy and
XX	localised neuropathies and central nervous system diseases, such as
XX	Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX	lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX	utilisation of the activities such as: Immune system suppression,
XX	CC Activin/inhibin activity, chemotactic/chemokinetic activity, hemostatic
XX	and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX	CC assays for receptor activity, arthritis and inflammation, leukaemia and
XX	CC C.N.S disorders.
XX	CC Note: The sequence data for this patent did not form part of the printed
XX	specification.
XX	Sequence 335 AA;
XX	Sequence 335 AA;
XX	Alignment Scores:
XX	Pred. No.: 3,34e-66 Length: 335
XX	Score: 914.00 Matches: 203
XX	Percent Similarity: 72.36% Conservative: 30
XX	Best Local Similarity: 63.04% Mismatches: 57
XX	Query Match: 30.81% Indels: 32
XX	DB: 22 Gaps: 8
XX	US-09-896-522-1 (1-1624) x AA140480 (1-335)
XX	CGCGCTCCGACCTCGGCGCTGGCGGCGCGCGCGCGCGGAGAGCGGCGCGCGGAGA 67
XX	8 CGCGCTCCGACCTCGGCGCTGGCGGCGCGCGCGCGCGGAGAGCGGCGCGCGGAGA 67
XX	Db 31 ArgThrArgProSerAlaProAla--AlaProSerValGlyCysGlyLysAlaArgArgGlu 49
XX	68 CCGAGTCCCGGGAGCGGAGCGCGCGAGATGCGCTTCGGCG-----G 106
XX	Db 50 SerAspAla-GlyAlaGlyGly-GluArgAlaSerValArgThrGlySerGlyArgArg 69
XX	107 GAGGCGGAAGACTCGAGAGACCCCGCGCGCGAGCGCGAGCGTCCGACGAGCG----- 158
XX	Db 69 LysGlyArgThrMetAlaGlyApsSerGluGlnThrLeuGlnAsnHisGlnInProAsn 89
XX	160 -----CCCTTCTCGATAGGAGGTAGCGCGGCACTGCGAGCGGGAAGTGCACCGTGT 211

Accession	Protein	Length	Score	E-value	Identity	Positives	Gaps	Conserved Domains
D8	1yGlycineProtein	1099	1099	1099	1099	1099	1099	1099
QY	212 GTGAMAGATCATGGAGTGTCTGGAGACGAGCGAGCGGACGAGAGTGG	271	271	271	271	271	271	271
Db	109 ysalalalytlvalalnlneuleuglylinalserglvalasphyryrarglnlysglnvaly	129	129	129	129	129	129	129
QY	212 TCATCTGAGCCAGGACAGATTCTACAAAGTCTTGAACGCGACAGACGAAGCCAGGCT	311	311	311	311	311	311	311
Db	129 alileuenserclnasperphenetyrargvalleuthrsergluglnlysalalalyalal	149	149	149	149	149	149	149
QY	332 TGAAGGACAGACAAATTTTGAACATCCAGAGCGCTTGATTAATGATTTGATCAGACA	390	390	390	390	390	390	390
Db	149 eulyysglylphneaspheasphsaproaspalaphneaspanglnuleilleleuylst	169	169	169	169	169	169	169
QY	392 CTCTGAMAAACATCGTGGAGGGCAAAACGCTGAGTGGCGGACCTTGAATTTTGTGACAC	451	451	451	451	451	451	451
Db	169 hrlleuylglulilehrtrgluglylstrhrvalglmlleprovaltyrarphevalserh	189	189	189	189	189	189	189
QY	452 ACTCAAGTTACACAGAGACCAACGAGTGTCTACCTCGGAGCGTGTCTGTGTTGAGGCA	511	511	511	511	511	511	511
Db	189 lsserarglygluglnhrvaltthrvaltyrprolaphalvalleuphegluglyl	209	209	209	209	209	209	209
QY	512 TCTTGATGTCTTACAGCCAGGAG--ATCCGGACATGTTCCACCTGGCCCTTTGTGTG	568	568	568	568	568	568	568
Db	209 lslleualaphetyserglngluarglilearapspheuglnmetlylsleuphevala	229	229	229	229	229	229	229
QY	569 ACACCGACCTCGACGTCAAGCTGTCTCGAAGATTCTCCGAGACG--CGCCGAGGGA	625	625	625	625	625	625	625
Db	229 spthrasphalaphsphrargleuenserargvalleuylasphllesergluarglyla	249	249	249	249	249	249	249
QY	626 GGGACCTTGAGACAGATTCTGACGACAGTACAC--ACCTTGATGAAGCCGGCTTGAGG	682	682	682	682	682	682	682
Db	249 rgaaspleuglucglnlleuenseraserthrleuargphevalylsprolaphagluc	269	269	269	269	269	269	269
QY	683 AGTTCTGCTGCTCCGACAAAGATATGCCGATGTATCATCCACGAGAGTGCACAAAT	742	742	742	742	742	742	742
Db	269 luphecyaleuproprolylstrylalaspvalilleleprokrglyalalaspasna	289	289	289	289	289	289	289
QY	743 TGGTGCATCAACCTGATCGTGCAGCATCATCCAGACATCTGGAATGGTGCATCTGCA	802	802	802	802	802	802	802
Db	289 rgvalproleasmeuilevalglmlsilelnhspilleuamnglyl-----	305	305	305	305	305	305	305
QY	803 AATGGACCGAGAGGATCCATATGGCGGAGCTACAGCGACCTTTTCTGACCGAGGG	862	862	862	862	862	862	862
Db	306 -----gylproserasnarglnthr-----AsnGlyc	315	315	315	315	315	315	315
QY	863 ACCACCTTGAGATGCTGACCTCTGCGAAACGGCTCACTTTGAGTCCAGACGACCC	922	922	922	922	922	922	922
Db	315 yslleuasnglyltyrthrproserarglylsarglnalasergluberserargrproh	335	335	335	335	335	335	335
QY	923 AC 924							
Db	335 ls 335							
RESULT 14								
ID ABB62307 standard; Protein; 260 AA.								
XX	AC	ABB62307;						
XX	AC	ABB62307;						
XX	DT	26-MAR-2002 (first entry)						
XX	DE	Drosophila melanogaster polypeptide SEQ ID NO 13713.						
XX	KX	Drosophila melanogaster polypeptide SEQ ID NO 13713.						
XX	KW	pharmaceutical.						
XX	XX	Drosophila melanogaster.						
XX	OS	WO200171042-A2.						
XX	XX	27-SEP-2001.						

XX 23-MAR-2001; 2001WO-US09231.
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li FMD, Myers EW;
 XX WPI, 2001-6656860/75.
 DR N-PSDB; ABL06410.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 PT
 XX
 PS Disclosure; SEQ ID NO 13713; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB16175), expressed DNA
 CC sequences (AB101840-AB16175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 260 AA;

Alignment Scores:
 Pred. No.: 7,846-52 Length: 260
 Score: 738.00 Matches: 143
 Percent Similarity: 82.94% Conservative: 32
 Best Local Similarity: 67.77% Mismatches: 34
 Query Match: 24.87% Indels: 2
 Gaps: 2

US-09-896-522-1 (1-1624) x ABB62307 (1-260)

QY 160 CCCTTCCTGATAGGGGTGAGCGCGGCGACTCCAGCGGAAGTGCACCTGTGTGAGAG 219
 Db 27 ProtheneuilegilyValaIaglyGlyThralaserilysserThrvalCyslyblys 46
 QY 220 ATCATGAGTGTCTGGAGCAGAACGAGGTGGAACGGCGGACCGGAGGTGTCATCTG 279
 Db 47 IlwecgluglnleuglylnalaglumerAsphisthrGlnArglnValalSerIle 66
 QY 280 AGCCAGAGCAGGTTCTCAAGGTCCTGACGCGAGAGCAAGGCCAGCTTGAAGA 339
 Db 67 SerGlnmsperPheIyrArgGluLeuThrProalagluysalalyalaglnlyblys 86
 QY 340 CAGTACATTTTGCATCCAGATGCTTGATATGATTTGATGCACAGACTCTGAG 399
 Db 87 LeuphensnheasphistProasphalPheasnclugluLeuetyrSerThrleuGln 106
 QY 400 AACATCGTGAGGGCAAAACGGTGAGGTGCCGACTGATATTTGTGACACACTCA 456
 Db 107 AsnilleuueyeglyHslyblyvalGluIleProserIyrAspIyrArgThrnsSerleu 126
 QY 457 AGGTTCACAGAGCAGCGGTGTCTACCTCGGACCTGGTCTGTGTGAGGCGATCTTG 516
 Db 127 AspPheGlnuansValleuValIleTyPrAlasPvalValleupheGluGlyIleLeu 146
 QY 517 GTGTTTACAGCCAGAGATCCGGGACATGTTCCACTCGCCCTCTGTGACACCGAC 576
 Db 147 ValPheIyrPheProIyIleArgGluLeuPheHsIleTySleupheValAspThrAsp 166
 QY 577 TCCGACGTCAGGCTGTCTCGAAAGTTCTCCGGAGCGTG--CGCCAGAGGAGGACCTG 633

Db 167 SerAspThrArgheulaArgArgValProArgAspIleasnGluArgGlyArgAspLeu 186
 QY 634 GAGCAGATTCTGACGAGTACACCACTTCGTGTAAGCCGCGCTTCGAGAGTCTCGCTG 693
 Db 187 AspAlaValleuThrGlnIyrMetThrPheVallybProalabheglugluPheCysSer 206
 QY 694 CCGACAAAGAGTATGCCGATGATCATCCACGAGAGTGCACATATGTTGCATC 753
 Db 207 ProThrlysbysPhehAlasPvalIleIleProArgGlyAlasPasnThrValAlaIle 226
 QY 754 AACCTGATCGTGACAGCATCCAGACATCTG 786
 Db 227 AspLeuIleValIleHsIleGlyIleLeu 237
 RESULT 15
 ABG27216
 ID ABG27216 standard; Protein; 326 AA.
 AC ABG27216;
 XX
 DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #27207.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX
 OS Homo sapiens.
 XX
 PN WO200175067-A2.
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Dmanac RT, Liu C, Tang YT;
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS91403.
 XX
 PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PT
 PS Claim 20; SEQ ID No 57575; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG0010-ABG3037 represent novel human
 CC diagnostic amino acid sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 326 AA:

Alignment Scores:

Pred. No.:	2.33e-41	Length:	326
Score:	610.50	Matches:	143
Percent Similarity:	59.52%	Conservative:	32
Best Local Similarity:	48.64%	Mismatches:	68
Query Match:	20.58%	Indels:	51
DB:	22	Gaps:	9

US-09-896-522-1 (1-1624) x ABG27216 (1-326)

```
QY 7 TCGCCTCCGACCTCGGCGCTGGGCGG-----CGC 36
Db 26 SerProGlySerGluArgLeuAlaGlyAspGlyHisValArgGluProArgGluArgLeu 45
QY 37 GCCGGGCCCCGGGAGAGGGCGGGCGGCGGACCGCATGC---GCGGAGCGGAGCGCGAG 93
Db 46 AlaSerAlaIaIaIaSerGlyGlySerProThrArgAlaArgAlaGlySerValArgPro 65
QY 94 ATGGCTTCGGCGGAGCGGAGAGCTGCGAGAGCCCGCGCG--GAGCGCAGCCGTCG 150
Db 66 PheAlaGlnAlaIaIaGlyGlyAlaAlaGluProTrpProGlyThrAlaSerArgPro 85
QY 151 CACGAG-----CGGCCCTTCCTGATAGGGGTGAGCGGCGC 186
Db 86 CysArgThrThrSerSerProThrAlaAlaSerThrSerLeuIleGlyValThrTrpGly 105
QY 187 ACTGCCAGCGGAGAGTGCACCGTGTGTGAGAGATCATGAGTGTCTGGAGCAAGAG 246
Db 106 ThrAlaSerGlyLysSerSerValCysAlaIaIaIeValGlnLeuLeuGlyGlnAnglu 125
QY 247 GTGGAACAGCGGAGCGGAGAGTGTCTATCTGAGCCAGAGAGTTTCAAGTCTCG 306
Db 126 ValAspTyrArgGlnArgGlnValIleLeuSerGlnAspSerPheTyrArgValLeu 145
QY 307 ACGGCGAGAGAGAGCGGAGCGCTTGAAGAGACAGTAAATTTTGACCATTCAGATGCC 366
Db 146 ThrSerGlnGlnIlySalAlaIaIaLeuLysGlyGlnPheAsnPheAspHisProAspAla 165
QY 367 TTTGATATGATTTTGATGCACAGAGACTCTGAGAAACATCGTGAAGGCGCAAAACGTGAG 426
Db 166 PheAspAsnGlnLeuIleLeuLysThrLeuLysGlnIleThrGlnGlyLysThrValGln 185
QY 427 GTGCCCACTATGATTTTGACACACTCAAGGTTACAGAGACCACGCTGTACCT 486
Db 186 IleProValTyrAspPheValSerHisSerArgLysGlnGluThrValThrValTyrPro 205
QY 487 GCGGAGCTGTGTTGTTGAGGGGACTTGTGTCTACAGCCAGAGATCCGGGACATG 546
Db 206 AlaAspAlaValLeuPheGlnGlyIleLeuAlaPheTyrSerGlnGlnValArgAspLeu 225
QY 547 TTCACCTGCGCTCTCTGTGACACCGACTCCGACGTCCAGGCTGTCTCGAAGA--GTT 603
Db 226 PheGlnMetLysLeuPheValaIaPheThrAspAlaAspThrArgLeuSerArgArgGlyIle 245
QY 604 CTCGGGAGCTGCGCGGAGGAGGAGCTTGAGCAATT-----642
Db 246 MetAsnLeuLeuHisProArgGlyLeuArgAlaIleThrIleAlaValPheGlyLys 265
QY 643 -----CTGACGCAGTACACCACTTCCTGGAAGCGCGCC-----675
Db 266 GlnAsnThrTyrIleArgLeuGlnProPheArgIleAsnValProProThrIleThrLys 285
QY 676 ---TTCGAGAGGATTC---TGCCTGCCG-----696
Db 286 HisIleGlnGlnLeuGlnCysAlaProValValThrGlnLeuSerArgLysGlnGluThr 305
QY 697 ---ACAAAGAGATGCGGATGTGATCATCCACGAGAGATG 735
Db 306 ValThrValTyrProAlaAspValValLeuPheGlnGlyIle 319
```

Search completed: November 25, 2003, 07:30:34
Job time : 86.9182 secs

This Page Blank (uspto)